

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 751 375**

51 Int. Cl.:

**G10L 15/18** (2013.01)

**G06F 17/27** (2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **25.02.2014** **PCT/EP2014/053643**

87 Fecha y número de publicación internacional: **04.09.2014** **WO14131763**

96 Fecha de presentación y número de la solicitud europea: **25.02.2014** **E 14708224 (2)**

97 Fecha y número de publicación de la concesión europea: **31.07.2019** **EP 2962296**

54 Título: **Análisis lingüístico basado en una selección de palabras y dispositivo de análisis lingüístico**

30 Prioridad:

**26.02.2013 DE 102013101871**

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

**31.03.2020**

73 Titular/es:

**PRECIRE TECHNOLOGIES GMBH (100.0%)**

**Charlottenburger Allee 40**

**52068 Aachen, DE**

72 Inventor/es:

**GRATZEL, DIRK C. y**

**GREB, CHRISTIAN**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

ES 2 751 375 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

Análisis lingüístico basado en una selección de palabras y dispositivo de análisis lingüístico

La invención se refiere a un procedimiento para el análisis lingüístico automatizado basado en una selección de palabras y a un dispositivo de análisis lingüístico.

5 La psicolingüística o psicología lingüística se refiere a los procesos de la adquisición lingüística, de la comprensión lingüística (conocimiento lingüístico) y de la producción lingüística:

- Adquisición lingüística: ¿cómo los niños pueden adquirir y aplicar conocimientos lingüísticos?
- Comprensión lingüística: ¿de qué manera y con qué conocimiento se forman y se crean los significados y el sentido de palabras, frases y textos?
- 10 - producción lingüística: ¿cómo se posibilitan las articulaciones lingüísticas con sentido?

Los psicólogos lingüísticas diferencian entre cuatro fases sucesivas:

- (1) la conceptualización de una idea, un pensamiento o un sentimiento que se desean expresar
- (2) la creación de un plan (de secuencia) lingüístico
- (3) la articulación como realización del plan
- 15 (4) la supervisión de la articulación

Estos procesos se dedican a los aspectos cognitivos de la lengua, referidos al pensamiento y a los conocimientos. A los psicólogos, en cambio, durante mucho tiempo apenas les interesaba la cuestión de cómo la personalidad se expresa en la lengua – salvo Freud, cuyas tesis, sin embargo, quedaron infructuosas empíricamente.

20 Además de modelos de análisis de calidad / contenido subjetivos, referidos al contexto, que en parte requieren un volumen extremadamente grande de tiempo y recursos, que se orientaban por unidades lingüísticas más grandes, como frases y textos enteros, en los años 1960 y 1970, con la aparición del ordenador, surgieron procedimientos de análisis objetivos de cantidad / forma automatizados, basados en unidades lingüísticas pequeñas, los llamados “métodos de conteo de palabras”.

25 El documento US5,696,981B1 describe un procedimiento y un dispositivo de análisis lingüístico para el análisis de personalización automatizado basado en una selección de palabras. Un programa almacenado en un ordenador personal contiene listas con palabras clave, comprendiendo cada lista una multiplicidad de palabras que están vinculados con uno de seis tipos de personalidad. Para el análisis de personalidad se registra un archivo de voz que comprende una multiplicidad de palabras de una persona que ha de ser analizada. Las palabras pueden proceder por ejemplo de un cuestionario, un monólogo o de un texto redactado por la persona. Se introducen como texto en el sistema informático o se convierten automáticamente en un archivo de voz por reconocimiento de voz durante un monólogo. En el marco del análisis de personalidad automatizado se examina si el archivo de voz contiene palabras clave de las seis listas con palabras clave. En un procedimiento de múltiples etapas, las palabras clave encontradas en el archivo de voz se ponderan y se asignan, en función del contexto, a base de las palabras clave contenidas en las seis listas, a uno de los seis tipos de personalidad. Para la emisión del resultado del análisis de personalidad, las palabras clave ponderadas se suman para cada uno de los seis tipos de personalidad y se realiza una asignación porcentual de la personalidad a los seis tipos de personalidad, que se representa especialmente en un diagrama de barras.

35 Una desventaja del procedimiento conocido consiste en que es posible sólo un análisis de personalidad automatizado basado en una selección de palabras. Otras características de la persona de la que procede el archivo de voz no se pueden determinar con este procedimiento. Las listas de palabras clave que tienen un reducido volumen se solapan sólo en una parte muy pequeña con los archivos de voz registrados. Por consiguiente, en el análisis de personalidad se tiene en consideración sólo una pequeña parte del archivo de voz. Se añade que las listas de palabras clave y la evaluación referida a estas son estáticas y, por tanto, la calidad del análisis de personalidad depende en gran medida de la selección de las palabras clave.

45 Por FRANCOIS MAIRESSE ET AL: “Using Linguistic Cues for the Automatic Recognition of Personality on Conversation and Text”, JOURNAL OF ARTIFICIAL INTELLIGENCE RESEARCH, tomo 30,

1 de noviembre de 2007 (2007-11-01), páginas 457 a 500 es un procedimiento asistido por ordenador para la determinación automática de la personalidad de una persona de ensayo por medio de archivos de voz, que comprende los pasos:

- en un primer paso se registran archivos de texto individuales de personas de ensayo.
- 50 - en un segundo paso, valoraciones de personalidad de personas de ensayo se vinculan con los archivos de texto.
- en un tercer paso, mediante conteo se determina la frecuencia de palabras en los distintos archivos de texto en 88 categorías.

- en un cuarto paso, por medio de métodos estadísticos se establecen relaciones entre las frecuencias contadas y las 88 categorías ("linguistic features" / características lingüísticas).
- en un quinto paso, el modelo se comprueba de forma asistida por ordenador y totalmente automática con la ayuda de personas adicionales, cuyas características de personalidad no son conocidas.

5 De JENNIFER GOLBECK ET AL: "Predicting personality with social media", PROCEEDINGS OF THE 2011 EXTENDED ABSTRACTS ON HUMAN FACTORS IN COMPUTING SYSTEMS, 1 de enero de 2011 (2011-01-01), páginas 253 a 262 y JENNIFER GOLBECK ET AL.: "Predicting personality from Twitter", PRIVACY SECURITY, RISK AND Trust (PASSAT), 2011 IEEE THIRD INTERNATIONAL CONFERENCE ON AND 2011 IEEE THIRD INTERNATIONAL CONFERENCE ON SOCIAL COMPUTING (SOCIALCOM), IEEE, 9 de octubre de 2011 (2011-10-09), páginas 149 a 156 se conocen además algoritmos de clasificación de personalidad, basados en el método Word Count (LIWC / conteo de palabras).

15 Partiendo de este estado de la técnica, la invención tiene el objetivo de proporcionar un procedimiento para el análisis automatizado de características sustancialmente discrecionales de una persona de la que procede un archivo de voz que ha de ser analizado, permitiendo el procedimiento una emisión de resultados intermedios aún durante el registro del archivo de voz. Además, se pretende proporcionar un dispositivo de análisis lingüístico.

Este objetivo se consigue mediante un procedimiento con las características de la reivindicación 1 y un dispositivo de análisis lingüístico con las características de la reivindicación 8.

20 El procedimiento y el dispositivo de análisis lingüístico permiten registrar de forma automatizada mensajes de voz y hablados por teléfono, teléfono móvil, correo electrónico, SMS, chat o blog, evaluarlos y deducir directamente de ellos características individuales / referidas a la personalidad, por ejemplo para un entrenamiento o un asesoramiento.

El dispositivo de análisis lingüístico permite deducir de forma extensa y fiable por ejemplo la personalidad y/o el estado psíquico del orador, y de esta manera, permitir por ejemplo a asesores, entrenadores o responsables de personal una comunicación más individual.

25 Para el análisis de características sustancialmente discrecionales de una persona, la invención se desprende del concepto conocido de la evaluación de listas de palabras clave estáticas referidas al tipo de personalidad. El análisis de características requiere una comparación con otras personas. Para ello, el procedimiento según la invención comprende una preparación del sistema informático según las características aa) a ad: mediante el almacenamiento de varios, por ejemplo al menos 1.000 archivos de voz de diferentes personas en una unidad de memoria del sistema informático según la característica aa) se forma una prueba al azar de referencia suficientemente amplia que permite la comparación con otras personas, necesaria para un reconocimiento de características, comprendiendo cada archivo de voz al menos una cantidad mínima de palabras, preferentemente al menos 100 palabras.

35 Cada archivo de voz y/o cada archivo de voz de referencia puede estar presente directamente como archivo de texto o como archivo de audio que mediante una transcripción se convierte en un archivo de texto. La entrada del archivo de texto se realiza a través de una unidad de entrada, por ejemplo un teclado. La conversión del archivo de audio se realiza especialmente por medio de un software para el reconocimiento de voz. A partir del archivo de audio, en una forma de realización del procedimiento según la invención, del archivo de voz puede extraerse adicionalmente información prosódica.

40 El archivo de diccionario almacenado según la característica ab) cubre al menos 80%, preferentemente sin embargo incluso al menos 95% de las palabras contenidas en los archivos de voz. Mediante archivos de diccionario tan amplios que comprenden al menos el vocabulario central, en lugar de usar listas de palabras clave, se garantiza que en el análisis se pueden tener en consideración grandes partes de los archivos de voz. Las palabras del archivo de diccionario se almacenan en una multiplicidad de diferentes categorías, estando clasificadas todas las palabras del archivo de diccionario en al menos una de las categorías. Las distintas categorías contienen por ejemplo diferentes clases de palabra tales como adverbios, pronombres, preposiciones, conjugaciones o pronombres indefinidos o palabras de un determinado entorno semántico, como por ejemplo negaciones o palabras relativas a emociones positivas o negativas. Las categorías también pueden contener combinaciones de palabras que se producen frecuentemente en un orden determinado, los llamados n-gramas. En total, están previstas preferentemente al menos 250 categorías de este tipo.

50 Mediante la comparación individual según la característica ac), se determina la frecuencia porcentual de las palabras contenidas en cada categoría del archivo de diccionario, con respecto a la cantidad de palabras de cada archivo de voz de referencia. Los valores determinados para la frecuencia porcentual se ponen según la característica ad), con la ayuda de métodos estadísticos y/o algorítmicos, en relación con las características conocidas de las personas de las que proceden los archivos de voz de referencia. De las relaciones se pueden deducir reglas que en un siguiente paso permiten predecir, con respecto a archivos de voz que no son parte de los archivos de voz de referencia registrados previamente y que proceden de personas de las que no se conocen características, características que previsiblemente presentan las personas de las que proceden los archivos de voz.

- Al crear el conjunto de reglas se origina un espacio tridimensional con tantas dimensiones como hay categorías

del archivo de diccionario y, dado el caso, información que ha de ser considerada según las reivindicaciones 8, 9 o 10. Los archivos de voz de referencia de la prueba al azar de referencia están dispuestos en dicho espacio en función de las frecuencias porcentuales en las distintas categorías y, dado el caso, de las propiedades de la información adicional.

- 5 - Con la ayuda de las características conocidas de las personas de las que proceden los archivos de voz de referencia, se forman grupos en este espacio multidimensional. Los archivos de voz de referencia de personas con una característica conocida, por ejemplo una alta escrupulosidad, se encuentran en determinadas partes del espacio multidimensional y los archivos de voz de referencia de personas con una baja escrupulosidad se encuentran en otras partes del espacio tridimensional.
- 10 - Con la ayuda de los métodos estadísticos y/o algorítmicos se separan de la forma más nítida posible las diferentes partes del espacio. Mediante la separación nítida, se garantiza que los archivos de voz registrados a continuación que no son parte de los archivos de voz de referencia, pueden asignarse con la ayuda del conjunto de reglas de la forma más exacta posible a una parte del espacio multidimensional.

A continuación de la preparación del sistema informático según las características aa) a ad), adicionalmente a los archivos de voz de referencia de la prueba al azar de referencia, se registran los archivos de voz de personas que no presentan características conocidas. Estos archivos de voz se analizan según la característica c):

- La comparación individual (característica ca)) del archivo de voz con el archivo de diccionario mediante la determinación de la frecuencia porcentual de las palabras contenidas en cada categoría del archivo de diccionario, con respecto a la cantidad de palabras del archivo de voz es requisito para el siguiente procesamiento del archivo de voz con la ayuda del conjunto de reglas según la característica cb) de la siguiente manera:
- El procesamiento de las frecuencias porcentuales, determinadas según la característica ca), en cada categoría se realiza con métodos estadísticos y/o algorítmicos. Las frecuencias porcentuales determinadas se examinan en cuanto a similitudes con las frecuencias porcentuales determinadas según la característica ac). El archivo de voz se clasifica conforme a las similitudes detectadas y se asigna a la propiedad de al menos una característica conocida que presentan las distintas personas de las que proceden los archivos de voz de referencia. Para mejorar la precisión del análisis, el examen de las similitudes puede limitarse a aquellas categorías del archivo de diccionario que son significativas para una característica determinada.

Las características asignadas a un archivo de voz se emiten en forma de un archivo de salida. El archivo de salida puede visualizarse directamente en una pantalla y/o reproducirse a través de un altavoz y/o imprimirse a través de una impresora. Sin embargo, el archivo de salida también puede suministrarse a procesos de procesamiento de datos adicionales, por ejemplo un proceso de transferencia de datos, con el que el archivo de salida se transfiere, a través de una red, a un ordenador receptor del sistema informático. Además, el archivo de salida puede suministrarse por ejemplo a un proceso de contestación automático que en función de la o las característica(s) asignada(s) al archivo de voz genera automáticamente una contestación. Las contestaciones pueden estar depositadas como contestaciones estándar en una base de datos que contiene asignaciones entre las características y las contestaciones estándar. La contestación generada automáticamente se transmite por ejemplo a la persona de la que procede el archivo de voz analizado. Sin embargo, la contestación también puede transmitirse también o adicionalmente a terceros como por ejemplo el explotador del análisis lingüístico automático basado en una selección de palabras. La contestación puede transmitirse como mensaje de texto y/o como archivo de audio. El mensaje de texto y/o el archivo de audio pueden reproducirse con una computadora del sistema informático, que presenta una unidad de visualización y un convertidor electroacústico, como por ejemplo un teléfono móvil o un ordenador personal.

Según las propiedades caracterizadoras de la invención, durante el registro del archivo de voz adicional, al menos una vez un archivo parcial del archivo de voz se almacena de forma intermedia en la unidad de memoria del sistema informático y se analiza el archivo parcial. El almacenamiento intermedio y el análisis permiten la visualización de resultados intermedios aún durante el registro del archivo de voz. De esta manera, se hace posible prácticamente un "análisis online" que puede ser de especial interés por ejemplo en el análisis de archivos de voz transmitidos a través del teléfono. Por ejemplo, en el trabajo policial pueden determinarse características de un llamante aún durante la conversación telefónica. El almacenamiento intermedio de un archivo parcial se realiza automáticamente o bien al alcanzar una cantidad determinada de palabras de un archivo de voz o tras transcurrir un tiempo determinado del registro del archivo de voz.

El archivo de salida y/o el archivo de salida provisional contienen preferentemente información relativa a características de personalidad y/o características que describen el estado psíquico de la persona de la que procede el archivo de voz. Para ello, la examinación de las similitudes del archivo de voz se limita a aquellas categorías del archivo de diccionario que son significativas para las características mencionadas.

- 55 Mediante las características de la reivindicación 2 se proporciona un procedimiento en el que está ampliada la base de datos de la muestra al azar de referencia. La ampliación de la base de datos mejora el poder informativo del análisis, ya que las relaciones estadísticas y/o algorítmicas del conjunto de reglas se apoyan en una muestra al azar de referencia más extensa. Para crear un procedimiento de autoaprendizaje está previsto un acuse de recibo que permite

una valoración de la exactitud de las características asignadas por vía del análisis del archivo de voz según la característica c). El acuse de recibo puede realizarse por ejemplo a través de una entrada en una pantalla de un ordenador personal, realizándose la valoración a través de campos de selección predefinidos como por ejemplo "procede, procede en parte, no procede". El acuse de recibo puede ser introducido por ejemplo por la persona de la que procede el archivo de voz. Si se amplía la muestra al azar de referencia según la característica f), el conjunto de reglas debe actualizarse teniendo en consideración la base de datos ampliada y volver a almacenarse, conforme a la característica ad). La actualización y el almacenamiento pueden realizarse en intervalos de tiempo después de haberse añadido a la muestra al azar de referencia varios archivos de voz.

Por el hecho de que a la muestra al azar de referencia se añaden sólo archivos de voz que presenten una cantidad mínima fijada de palabras, se garantiza que se añaden exclusivamente archivos de voz de la muestra al azar de referencia que son adecuadas para la mejora del conjunto de reglas. La cantidad mínima de palabras asciende preferentemente a 100 palabras.

Para hacer posible por ejemplo una preselección del idioma o una mejor adaptación a las características de la persona de la que procede el archivo de voz, que han de ser reconocidas por el procedimiento, en una forma de realización de la invención se propone que el archivo de diccionario se almacene en el sistema informático en función del uso previsto del procedimiento. Preferentemente, según la reivindicación 7, en el sistema informático están almacenados varios archivos de diccionario con diferentes contenidos, que pueden activarse opcionalmente en función del uso previsto del procedimiento.

En la forma de realización del procedimiento según la invención, conforme a la reivindicación 7, se sigue mejorando el resultado del análisis lingüístico automatizado basado en una selección de palabras. Adicionalmente, se registra al menos una información adicional para cada archivo de voz de referencia y para cada archivo de voz que ha de ser analizado. En la información adicional se trata especialmente de información prosódica, como por ejemplo la duración (p.ej., velocidad del habla, pausas), la frecuencia (p.ej., tono medio, tono máximo) y la energía (p.ej., el volumen en dB) o información sintáctica y/o morfológica. En concreto, la consideración de la información adicional en el marco del análisis lingüístico basado en una selección de palabras requiere las siguientes medidas:

- la preparación del sistema informático según la característica a) se amplía por el paso de que adicionalmente se registra la al menos una información adicional de cada archivo de voz de referencia. Esta información, por ejemplo, el volumen, la duración o la frecuencia, están presentes como valor de medición para cada archivo de voz de referencia y, por tanto, puede ser procesada directamente por el conjunto de reglas. También una información sintáctica y/o morfológica, por ejemplo, la frecuencia del uso de sustantivos, está presente como valor numérico para cada archivo de voz de referencia y, por tanto, puede ser procesada directamente por el conjunto de reglas.
- el conjunto de reglas está concebido para el procesamiento directo de la información adicional, de tal forma que con métodos estadísticos y/o algorítmicos determina además relaciones entre la información adicional, registrada como valores de medición o numéricos, y las características conocidas de las personas de las que proceden los archivos de voz de referencia.
- el registro y almacenamiento subsiguientes de un archivo de voz que ha de ser analizado, según la característica b), requiere adicionalmente el registro de la al menos una información adicional relativa a cada archivo de voz.
- el análisis del archivo de voz según la característica c) comprende adicionalmente al procesamiento de las frecuencias porcentuales determinadas según la característica ca) con la ayuda del conjunto de reglas, un procesamiento de la al menos una información adicional de cada archivo de voz registrado, examinando el conjunto de reglas con los métodos estadísticos y/o algorítmicos la al menos una información adicional de cada archivo de voz en cuanto a similitudes con dicha al menos una información adicional en los archivos de voz de referencia. El conjunto de reglas clasifica el archivo de voz implicando todas las similitudes detectadas con los archivos de voz de referencia en el espacio multidimensional definido por las categorías y por cada información adicional y lo asigna a la propiedad de al menos una característica conocida que presentan las diferentes personas de las que proceden los archivos de voz de referencia.

Como ya se ha explicado al principio, el archivo de salida puede suministrarse a un proceso de contestación automático que en función de la(s) característica(s) asignada(s) al archivo de voz genera automáticamente una contestación. El archivo de contestación se crea como archivo de audio. La reproducción del archivo de video se realiza por medio de un convertidor electroacústico conectado al sistema informático. Por el hecho de que la reproducción del archivo de audio con el convertidor electroacústico es controlada en función de las características contenidas en el archivo de salida, la reproducción de la contestación puede adaptarse por ejemplo a la personalidad y/o al estado psíquico de la persona de la que procede el archivo de voz analizado que ha de ser contestado. La reproducción del archivo de audio es controlado en función de las características contenidas en el archivo de salida, preferentemente de tal forma que se modifican la duración y/o la frecuencia y/o la energía de la reproducción.

La adaptación técnica de la reproducción del archivo de audio puede emplearse por ejemplo en un sistema de contestador automático. Si el análisis lingüístico basado en una selección de palabras detecta un llamante miedoso o neurótico, se ralentiza la reproducción de voz y se baja la frecuencia.

El dispositivo de análisis lingüístico para el análisis lingüístico automatizado basado en una selección de palabras comprende

- un sistema informático con al menos una unidad de memoria,
- una unidad de entrada conectada al sistema informático,
- 5 - una unidad de salida conectada al sistema informático y
- un programa que está configurado para la ejecución de un procedimiento según una de las reivindicaciones 1 a 12.

En el sentido de la invención, el sistema informático puede comprender sólo una computadora o varias computadoras interconectadas. Una computadora del sistema informático es un aparato que procesa datos por medio de un programa de procesamiento de datos. Por lo tanto, el sistema informático puede comprender tanto ordenadores personales usuales en el mercado, pero también smartphones y otros aparatos electrónicos programables. La conexión entre varias computadoras puede establecerse a través de una red o de varias redes, especialmente también a través de Internet. El sistema informático puede presentar una o varias unidades de memoria en las que se almacenan los distintos archivos tales como los archivos de diccionario y de voz así como los archivos de programa. Los archivos pueden almacenarse de forma distribuida entre varias unidades de memoria o todas en una sola unidad de memoria.

Si el sistema informático presenta varias unidades de memoria, los archivos de voz y los archivos de diccionario pueden almacenarse en diferentes unidades de memoria para permitir un acceso rápido a los distintos archivos para la comparación individual que ha de ser realizada por el programa.

Si el registro de los archivos de voz de referencia o de los archivos de voz no se realiza de forma basada en texto, por ejemplo por medio de un teclado, sino que la voz hablada debe registrarse como archivo de audio, la unidad de entrada presenta un sistema de detección de voz conocido para la transcripción del archivo de audio en un archivo de texto. Además, el sistema informático puede estar concebido para la extracción de información prosódica de cada archivo de audio.

Como unidad de salida entran en consideración una pantalla y/o a través de un altavoz y/o una impresora. En lugar de emitir el archivo de salida directamente a través de la unidad de salida, el archivo de salida puede suministrarse a través de una interfaz a un proceso de procesamiento de datos adicional y el resultado del proceso de procesamiento de datos se emite con la unidad de salida, por ejemplo un altavoz.

Por ejemplo, en función de las características contenidas en el archivo de salida (por ejemplo, una característica de personalidad de la persona) puede ser controlado automáticamente un dispositivo técnico, por ejemplo una central de teléfono, para vincular la persona de la que procede el archivo de voz analizado, automáticamente con una persona de contacto adecuada para su personalidad.

A continuación, con la ayuda de las figuras 1 a 9 se explican en detalle los pasos del procedimiento según la invención para el análisis lingüístico automatizado basado en una selección de palabras:

Las figuras 1 a 5 ilustran la preparación del sistema informático del dispositivo de análisis lingüístico: en primer lugar, como se puede ver en la figura 1, varios (p.ej. 1.000) archivos de voz de referencia 1 a n (1.10) se almacenan en una unidad de memoria (1.20) del sistema informático (1.30). Estos archivos de voz de referencia (1.10) forman una muestra al azar de referencia (1.40). Un archivo de voz de referencia (1.10) puede estar presente o bien directamente como archivo de texto, o bien, como archivo de audio hablado que mediante una transcripción con la ayuda de un sistema de reconocimiento de voz se convierte en un archivo de texto. Cada archivo de voz de referencia (1.10) comprende una cantidad mínima de por ejemplo al menos 100 palabras. Todos los archivos de voz de referencia (1.10) que han de ser almacenados en la muestra al azar de referencia (1.40) proceden de personas diferentes.

En la figura 2 se puede ver el almacenamiento de un archivo de diccionario (2.20) con una multiplicidad de distintas categorías 1 a n (2.10) en la unidad de memoria (1.20) del sistema informático (1.30). En las preferentemente más de 250 distintas categorías (2.10) se trata por ejemplo de las categorías: adverbios, pronombres, negaciones, artículos o palabras de un determinado entorno semántico, o de palabras que aparecen frecuentemente juntas, n-gramas etc. Todas las palabras del archivo de diccionario (2.20) están clasificadas en al menos una de las categorías (2.10).

La figura 3 ilustra los pasos para la preparación de un conjunto de reglas (5.40) (véase la figura 5) que para la preparación del sistema informático (1.3) se almacena en la unidad de memoria (1.20). Se registran los archivos de voz de referencia (1.10) que proceden respectivamente de diferentes personas de ensayo y se almacenan en la unidad de memoria (1.20). Estos presentan características (4.20) conocidas, como por ejemplo características de personalidad, motivos, estímulos conocidos o un estado psíquico conocido. Los archivos de voz de referencia (1.10), con características (4.20) conocidas, se analizan con la ayuda del archivo de diccionario (2.20). Durante ello se realizan comparaciones individuales de cada archivo de voz de referencia (1.10) de la muestra al azar de referencia (1.40) con el archivo de diccionario (2.20) depositado en la unidad de memoria (1.20). Se determinan frecuencias porcentuales (3.40) de las palabras de cada archivo de voz de referencia (1.10), contenidas en cada categoría (2.10) del archivo de diccionario (2.20), con respecto a la cantidad total de palabras de dicho archivo de

voz de referencia. Estas frecuencias porcentuales (3.40) se almacenan en una unidad de memoria (1.20) del sistema informático (1.30) para su siguiente procesamiento.

En la figura 4 está representada la extracción opcional de información adicional de los archivos de voz de referencia (1.10) con características (4.20) conocidas, según las reivindicaciones 8 a 10. Si el archivo de voz de referencia (1.10) se transcribe a partir de un archivo de audio, como información adicional puede obtenerse información (2.30) prosódica y almacenarse en la unidad de memoria (1.20) del sistema informático (1.30) para el siguiente procesamiento. Esto se realiza mediante el análisis del archivo de audio que ha de ser transcrito, con programas conocidos por el estado de la técnica para la extracción de características lingüísticas prosódicas de archivos de audio, como por ejemplo con el programa de fuente abierta "Praat" para análisis fonéticos a base de señales que fue desarrollado por Paul Boersma y David Weenink en el 'Institute of Phonetic Sciences' en la Universidad de Amsterdam y que se ha establecido como casi estándar en el campo de las ciencias de las lenguas y de la comunicación.

En el ejemplo de realización, como información prosódica, (2.30) se detecta la energía (por ejemplo, volumen en decibelios) del archivo de audio.

Además, en el ejemplo, adicionalmente, a partir de cada archivo de audio transcrito a partir del archivo de audio se extrae información morfológica (2.40) y sintáctica (2.50). La información morfológica (2.40) se refiere a las distintas posibilidades de formar palabras – se trata de una forma de "gramática léxica". Se trata de conjugación, declinación, combinación, derivación y abreviación de palabras que se encuentra en un archivo de voz de referencia (1.0). Se cuenta la cantidad de palabras con una determinada "gramática léxica". La sintaxis describe la "gramática léxica" en la que se examina el tipo de sintaxis. Esto se realiza con 'part-of-speech tagger' (marcadores de parte de la oración) definidos especialmente. También esta información está presente como valor numérico. La información se almacena, de forma análoga a las frecuencias porcentuales (3.40) en las categorías (2.10), en una unidad de memoria (1.20) del sistema informático (1.30) para su siguiente cálculo.

Las frecuencias porcentuales (3.40) de cada archivo de voz de referencia (1.10) en las distintas categorías (2.10) del archivo de diccionario (2.20) y los valores numéricos con respecto a la información extraída eventualmente adicionalmente, se procesan en el conjunto de reglas (5.40). La representación en la figura 5 muestra para mayor claridad únicamente cómo las frecuencias porcentuales (3.40) de cada archivo de voz de referencia (1.10) se reúnen con las características (4.20) conocidas de los mismos archivos de voz de referencia (1.10) en el conjunto de reglas (5.40) y se comparan.

En el conjunto de reglas (5.40), cada categoría (2.10) se dispone formando una dimensión de un espacio multidimensional (5.30). Según la frecuencia porcentual (3.40) de las palabras contenidas en cada categoría (2.10), los archivos de voz de referencia (1.10) de todas las personas con características (4.20) conocidas están dispuestos en el espacio multidimensional (5.30). La disposición se realiza con métodos estadísticos y/o algorítmicos, se determinan las relaciones al menos entre las frecuencias porcentuales (3.40) y las características conocidas de las personas de las que proceden los archivos de voz de referencia (1.10). Como método algorítmico de separación nítida se emplea por ejemplo un procedimiento conocido como "Support Vector Machine" (máquina de vector de soporte). En una forma de realización, este procedimiento tiene como objetivo tender un hiperplano (recta multidimensional) en el espacio multidimensional de tal forma que en las por ejemplo 100 dimensiones se separa lo mejor posible en el espacio entre personas con una alta, media y baja intensidad de una determinada característica conocida. Como método estadístico resultan apropiadas especialmente las regresiones multidimensionales. Independientemente del método estadístico o algorítmico empleado, se reconocen reglas de cómo están interrelacionados los archivos de voz de referencia y las características conocidas de las personas de las que proceden los archivos de voz de referencia. Esto permite entonces en un paso de análisis subsiguiente de un archivo de voz (6.10) de una persona con características desconocidas una predicción de las características de la persona de la que procede el archivo de voz (6.10) adicional (véase la figura 6). El conjunto de reglas (5.40) se almacena en la unidad de memoria (1.20) del sistema informático (1.30) (véase la figura 5).

A continuación de la preparación del sistema informático se realizan el registro y el almacenamiento, representados en la figura 6, del archivo de voz (6.10), cuyas características no son conocidas, adicionalmente a los archivos de voz (1.10) de la muestra al azar de referencia (1.40), en la unidad de memoria (1.20) del sistema informático (1.30). El archivo de voz (6.10) puede estar presente o bien directamente como archivo de texto, o bien, como archivo de audio hablado que con la ayuda de un sistema de reconocimiento de voz se convierte en un archivo de texto. Cada archivo de texto (6.10) comprende preferentemente una cantidad mínima de palabras, especialmente al menos 100 palabras.

La figura 7 ilustra el primer paso del análisis del archivo de voz (6.10) con la ayuda del archivo de diccionario (2.20) almacenado en la unidad de memoria (1.20). Se realizan comparaciones individuales del archivo de voz (6.10) con el archivo de diccionario (2.20) para determinar las frecuencias porcentuales (7.30) de las palabras en el archivo de voz (6.10) que están contenidas en cada categoría (2.10) del archivo de diccionario (2.20). Las frecuencias porcentuales (7.30), determinadas en el primer paso de análisis, del archivo de voz (6.10) se procesan en un segundo paso de análisis con la ayuda del conjunto de reglas (5.40) de tal forma que las frecuencias porcentuales (7.30) determinadas con métodos estadísticos y/o algorítmicos en el primer paso de análisis se examinan en cuanto a similitudes con las frecuencias porcentuales (3.40) determinadas durante la preparación del sistema informático, y el archivo de voz (6.10)

se clasifica conforme a las similitudes detectadas, es decir, se ordena en el espacio multidimensional (5.30) y se asigna a al menos una característica conocida que presentan las distintas personas de las que proceden los archivos de voz de referencia (1.10). En cuanto a los métodos estadísticos y/o algorítmicos que entran en consideración se remite a las indicaciones referentes a la figura 5.

- 5 Los resultados de este análisis (características (4.20) asignadas) se almacenan en un archivo de salida (8.20) en una unidad de memoria (1.20) del sistema informático (1.30) (véase la figura 8). La asignación puede indicarse por ejemplo como intensidad en % de la característica (4.20).

10 La figura 8 ilustra el procesamiento de las características (4.20) de un archivo de voz (6.10) asignadas con la ayuda del conjunto de reglas (5.40). El archivo de salida (8.20) contenido en las características (4.20) asignadas se emite a través de una unidad de salida (9.10, 9.20, 9.30, 9.40) del sistema informático. La unidad de salida puede comprender una unidad de visualización (9.10) y/o una impresora (9.20) y/o una interfaz (9.30) y/o un convertidor electroacústico (9.40).

15 La figura 9 ilustra una forma de realización del análisis lingüístico basado en una selección de palabras según la invención en el que para el archivo de salida (8.20) es creado automáticamente por un módulo de contestación (10.20) un archivo de contestación (10.10) como archivo de audio. Las contestaciones pueden estar depositadas como contestaciones estándar en una base de datos que contiene asignaciones entre las características (4.20) y las contestaciones estándar. El archivo de audio se reproduce por medio de un convertidor electroacústico (12.10). Un módulo de control (11.40) controla la reproducción del archivo de contestación (10.10) presente como archivo de audio, con el convertidor electroacústico (12.10), en función de las características contenidas en el archivo de salida (8.20).  
20 El módulo de control (11.40) controla la energía (11.10), la frecuencia (11.20) y la duración (11.30) de la reproducción. El convertidor electroacústico (12.10) convierte las señales eléctricas, generadas por el módulo de control (11.40), en señales acústicas. Por ejemplo, la duración de la reproducción del archivo de contestación se ralentiza (menor velocidad de habla) y la frecuencia de la reproducción se baja con respecto a los valores de partida predefinidos para la reproducción del archivo de audio, cuando de la característica (4.20) asignada se deduce que se trata de una  
25 persona especialmente nerviosa de la que procede el archivo de voz (6.10) al que está contestando automáticamente el módulo de contestación (10.10).

A diferencia de la representación en la figura 8, el archivo de salida (8.20) no se emite directamente, sino que, a través de una interfaz (9.30) se suministra a los demás procesos de procesamiento de datos por el módulo de contestación (10.20) y el módulo de control (11.40). El módulo de contestación (10.20) y/o el módulo de control (11.40) pueden  
30 estar almacenados o bien en la misma computadora del sistema informático (1.30) que el archivo de salida (8.20), o bien, en una computadora del sistema informático (1.30), que está acoplada a través de una red y que se encuentra por ejemplo en el lugar de un cliente.

#### Lista de signos de referencia

| Núm. | Designación                                             |
|------|---------------------------------------------------------|
| 1.10 | Archivos de voz de referencia                           |
| 1.20 | Unidad de memoria                                       |
| 1.30 | Sistema informático                                     |
| 1.40 | Muestra al azar de referencia                           |
|      |                                                         |
| 2.10 | Categorías                                              |
| 2.20 | Archivo de diccionario                                  |
| 2.30 | Información prosódica                                   |
| 2.40 | Información morfológica                                 |
| 2.50 | Información sintáctica                                  |
|      |                                                         |
| 3.40 | Frecuencia porcentual (preparación sistema informático) |

(continuación)

| Núm.  | Designación                                  |
|-------|----------------------------------------------|
|       |                                              |
| 4.20  | Características conocidas                    |
|       |                                              |
| 5.30  | Espacio multidimensional (espacio vectorial) |
| 5.40  | Conjunto de reglas                           |
| 6.10  | Archivo de voz                               |
|       |                                              |
| 7.30  | Frecuencia porcentual (análisis)             |
|       |                                              |
| 8.20  | Archivo de salida                            |
|       |                                              |
| 9.10  | Unidad de visualización                      |
| 9.20  | Impresora                                    |
| 9.30  | Interfaz                                     |
| 9.40  | Convertidor electroacústico                  |
| 10.10 | Archivo de contestación                      |
| 10.20 | Módulo de contestación                       |
|       |                                              |
| 11.10 | Energía                                      |
| 11.20 | Frecuencia                                   |
| 11.30 | Duración                                     |
| 11.40 | Módulo de control                            |
|       |                                              |
| 12.10 | Convertidor electroacústico                  |

## REIVINDICACIONES

1. Procedimiento para el análisis lingüístico automatizado basado en una selección de palabras, que comprende los pasos:

a) la preparación de un sistema informático (1.30) mediante

- 5 aa) el almacenamiento de varios archivos de voz de referencia (1.10) en una unidad de memoria (1.20) del sistema informático (1.30) para formar una muestra al azar de referencia (1.40), comprendiendo cada archivo de voz de referencia (1.10) una cantidad mínima de 100 palabras y procediendo todos los archivos de voz de referencia (1.10) de distintas personas que presentan características conocidas,
- 10 ab) el almacenamiento de un archivo de diccionario (2.20) con al menos 250 distintas categorías (2.10) en una unidad de memoria (1.20) del sistema informático (1.30), cubriendo el archivo de diccionario almacenado, que comprende el vocabulario central, al menos el 95 % de las palabras contenidas en los archivos de voz, y estando clasificadas todas las palabras del archivo de diccionario (2.20) en al menos una de las al menos 250 categorías (2.10),
- 15 ac) la realización de una comparación individual de cada archivo de voz de referencia (1.10) de la muestra al azar de referencia (1.40) con el archivo de diccionario (2.20) mediante la determinación de la frecuencia porcentual (3.40) de las palabras en cada archivo de voz de referencia (1.10) que están contenidas en cada categoría (2.10) del archivo de diccionario (2.20),
- 20 ad) el almacenamiento de un conjunto de reglas (5.40) en una unidad de memoria (1.20) del sistema informático (1.30), que con métodos estadísticos y/o algorítmicos determina relaciones al menos entre las frecuencias porcentuales (3.40), determinadas según la característica ac), en una o varias categorías (2.10), y al menos una característica (4.20) conocida de las personas de las que proceden los archivos de voz de referencia (1.10),

b) a continuación de la preparación del sistema informático según las características aa) a ad), el registro y el almacenamiento de un archivo de voz (6.10), adicionalmente a los archivos de voz de referencia (1.10) de la muestra al azar de referencia (1.40), en una unidad de memoria (1.20) del sistema informático (1.30), estando presentes cada archivo de voz (6.10) y cada archivo de voz de referencia directamente como archivo de texto o como archivo de audio, que mediante una transcripción se convierte en un archivo de texto,

- 25 c) el análisis del archivo de voz (6.10) registrado y almacenado adicionalmente según la característica b), mediante
  - ca) la realización de una comparación individual del archivo de voz (6.10) con el archivo de diccionario (2.20) mediante la determinación frecuencia porcentual (7.30) de las palabras en el archivo de voz (6.10) que están contenidas en cada categoría (2.10) del archivo de diccionario (2.20),
  - 30 cb) el procesamiento de las frecuencias porcentuales (7.30) determinadas según la característica ca), con la ayuda del conjunto de reglas (5.40) que con métodos estadísticos y/o algorítmicos examina las frecuencias porcentuales (7.30) determinadas según la característica ca) en cuanto a similitudes con las frecuencias porcentuales (3.40) determinadas según la característica ac), y que clasifica el archivo de voz (6.10) conforme a las similitudes detectadas y las asigna a al menos una característica conocida que presentan las personas de las que proceden los archivos de voz de referencia (1.10),

d) la creación de un archivo de salida (8.20) que contiene características (4.20) asignadas al archivo de voz (6.10) según la característica cb),

e) la emisión del archivo de salida (8.20),

40 **en el que**

ga) durante el registro del archivo de voz (6.10) adicional según la característica b), al menos una vez un archivo parcial del archivo de voz (6.10) se almacena de forma intermedia en la unidad de memoria (1.20) del sistema informático (1.30), comprendiendo cada archivo de voz (6.10) adicional al menos 100 palabras,

gb) el archivo parcial almacenado de forma intermedia se analiza mediante

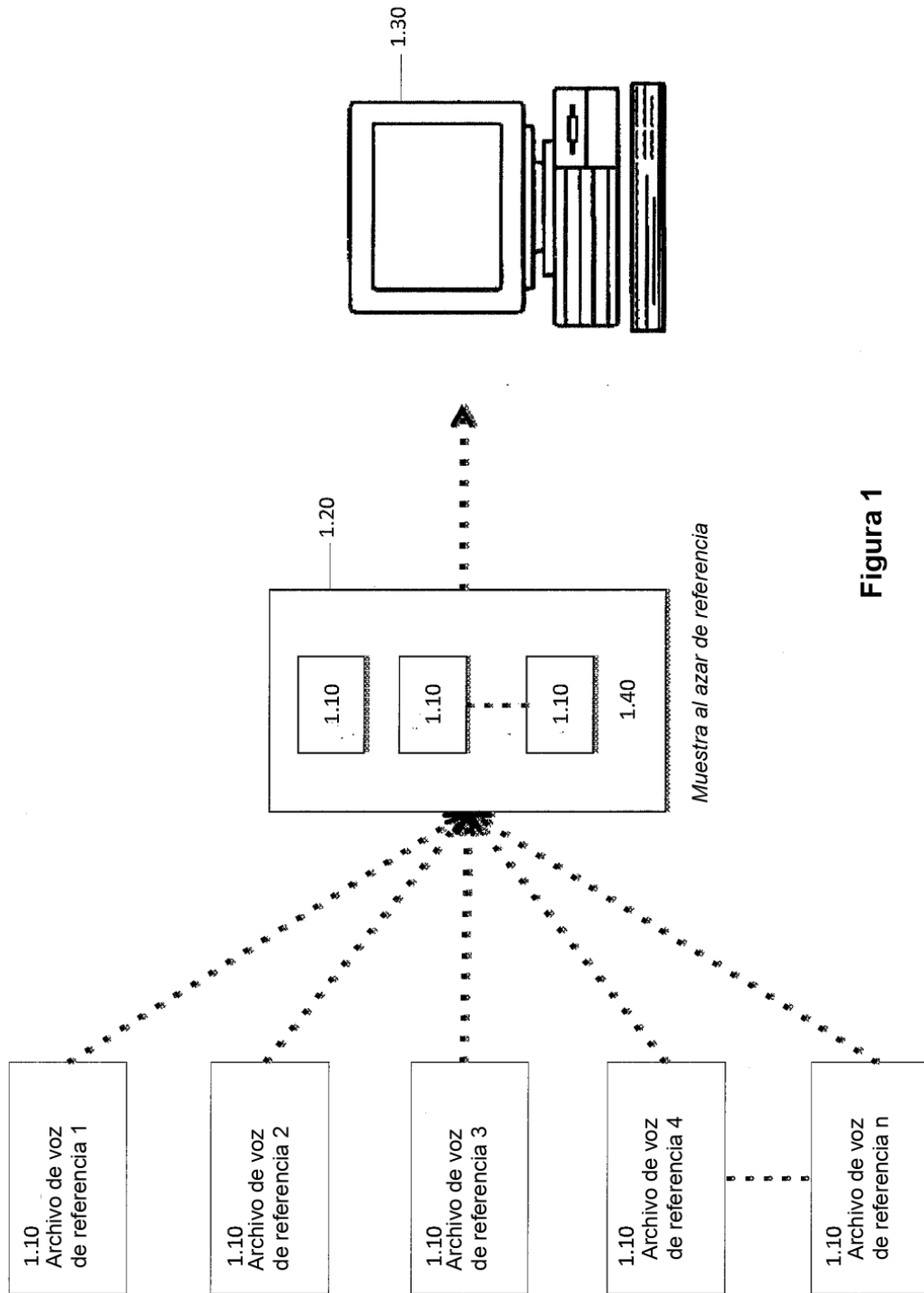
- 45 gba) la realización de una comparación individual del archivo parcial con el archivo de diccionario (2.20) mediante la determinación de la frecuencia porcentual (7.30) de las palabras en el archivo parcial que están contenidas en cada categoría (2.10) del archivo de diccionario (2.20),
- 50 gbb) el procesamiento de las frecuencias porcentuales (7.30) determinadas según la característica gba) con la ayuda del conjunto de reglas (5.40) que con métodos estadísticos y/o algorítmicos examina las frecuencias porcentuales (7.30) determinadas según la característica gba) en cuanto a similitudes con las frecuencias porcentuales (3.40) determinadas según la característica ac), y que clasifica el archivo parcial conforme a las similitudes detectadas y lo asigna a al menos una característica conocida que presentan las distintas personas de las que proceden los archivos de voz de referencia (1.10),

55 gc) la creación de un archivo de salida provisional que contiene características (4.20) asignadas al archivo parcial según la característica gbb),

gd) la emisión del archivo de salida provisional.

2. Procedimiento según la reivindicación 1, **caracterizado por** los pasos adicionales:

- fa) la ampliación de la muestra al azar de referencia (1.40) según la característica aa) mediante la adición según la característica b) de archivos de voz (6.10) adicionales, registrados según la característica b), como archivos de voz de referencia (1.10),
- fb) la previsión de un acuse de recibo mediante una entrada que permite una valoración de la exactitud de las características asignadas en el marco del análisis del archivo de voz según la característica c) y
- fc) la actualización y el nuevo almacenamiento del conjunto de reglas (5.40) teniendo en consideración la base de datos ampliada conforme a la característica ad).
3. Procedimiento según la reivindicación 2, **caracterizado porque** a la muestra al azar de referencia (1.40) se añaden sólo aquellos archivos de voz (6.10) adicionales que presenten una cantidad mínima fijada de al menos 100 palabras.
4. Procedimiento según una de las reivindicaciones 1 a 3, **caracterizado porque** el archivo de salida y/o el archivo de salida provisional contienen características de personalidad y/o características relativas al estado psíquico de la persona.
5. Procedimiento según una de las reivindicaciones 1 a 4, **caracterizado porque** en el sistema informático (1.30) se almacenan distintos archivos de diccionario (2.20) en función del uso previsto del procedimiento.
6. Procedimiento según una de las reivindicaciones 1 a 4, **caracterizado porque** en el sistema informático (1.30) están almacenados varios archivos de diccionario (2.20) de distinto contenido que se pueden usar opcionalmente.
7. Procedimiento según una de las reivindicaciones 1 a 6, **caracterizado porque**
- la preparación del sistema informático (1.30) según la característica a) comprende adicionalmente el registro y el almacenamiento de al menos una información (2.30) adicional de cada archivo de voz de referencia (1.10) y el conjunto de reglas (5.40) está concebido de tal forma que con métodos estadísticos y/o algorítmicos determina además relaciones entre la al menos una información (2.30) adicional y las características (4.20) conocidas de las personas de las que proceden los archivos de voz de referencia (1.10),
  - el registro y almacenamiento subsiguientes de un archivo de voz (6.10) según la característica b) comprende adicionalmente el registro y el almacenamiento de la al menos una información (2.30) adicional relativa a cada archivo de voz (6.10), y
  - el análisis del archivo de voz (6.10) según la característica c) comprende adicionalmente al procesamiento de las frecuencias porcentuales determinadas según la característica ca) con la ayuda del conjunto de reglas (5.40), un procesamiento de la al menos una información adicional de cada archivo de voz (6.10) registrado, en donde el conjunto de reglas (5.40) examina con los métodos estadísticos y/o algorítmicos la al menos una información (2.30) adicional de cada archivo de voz en cuanto a similitudes con esta al menos una información (2.30) adicional en los archivos de voz de referencia, y en donde el conjunto de reglas (5.40) examina el archivo de voz (6.10) implicando todas las similitudes detectadas y lo asigna a la propiedad de al menos una característica (4.20) conocida que presentan las diferentes personas de las que proceden los archivos de voz de referencia (1.10),
  - en donde archivo de voz de referencia y el archivo de voz (1.10, 6.10) se transcriben cada uno desde un archivo de audio y como información adicional se extrae información prosódica de los archivos de audio y/o como información adicional se extrae información morfológica y/o sintáctica (2.40, 2.50) a partir de cada archivo de voz de referencia y de cada archivo de voz (1.10, 6.10).
8. Dispositivo de análisis lingüístico para el análisis lingüístico automatizado basado en una selección de palabras, que comprende
- un sistema informático (1.30) con al menos una unidad de memoria (1.20),
  - una unidad de entrada conectada al sistema informático (1.30),
  - un programa almacenado en al menos una unidad de memoria (1.20) que está concebido para la realización de un procedimiento según una de las reivindicaciones 1 a 7, en el sistema informático (1.30).
9. Dispositivo de análisis lingüístico según la reivindicación 8, **caracterizado porque** el sistema informático (1.30) presenta varias unidades de memoria (1.20) y los archivos de voz (1.10, 6.10) y los archivos de diccionario (2.20) están almacenados en diferentes unidades de memoria (1.20).
10. Dispositivo de análisis lingüístico según las reivindicaciones 8 o 9, **caracterizado porque** la unidad de entrada presenta un sistema de reconocimiento de voz.
11. Dispositivo de análisis lingüístico según una de las reivindicaciones 8 a 10, **caracterizado porque** al sistema informático están conectados una impresora (9.20) y/o una unidad de visualización (9.10) y/o un convertidor electroacústico (9.40) para reproducir el archivo de salida (8.20).



**Figura 1**

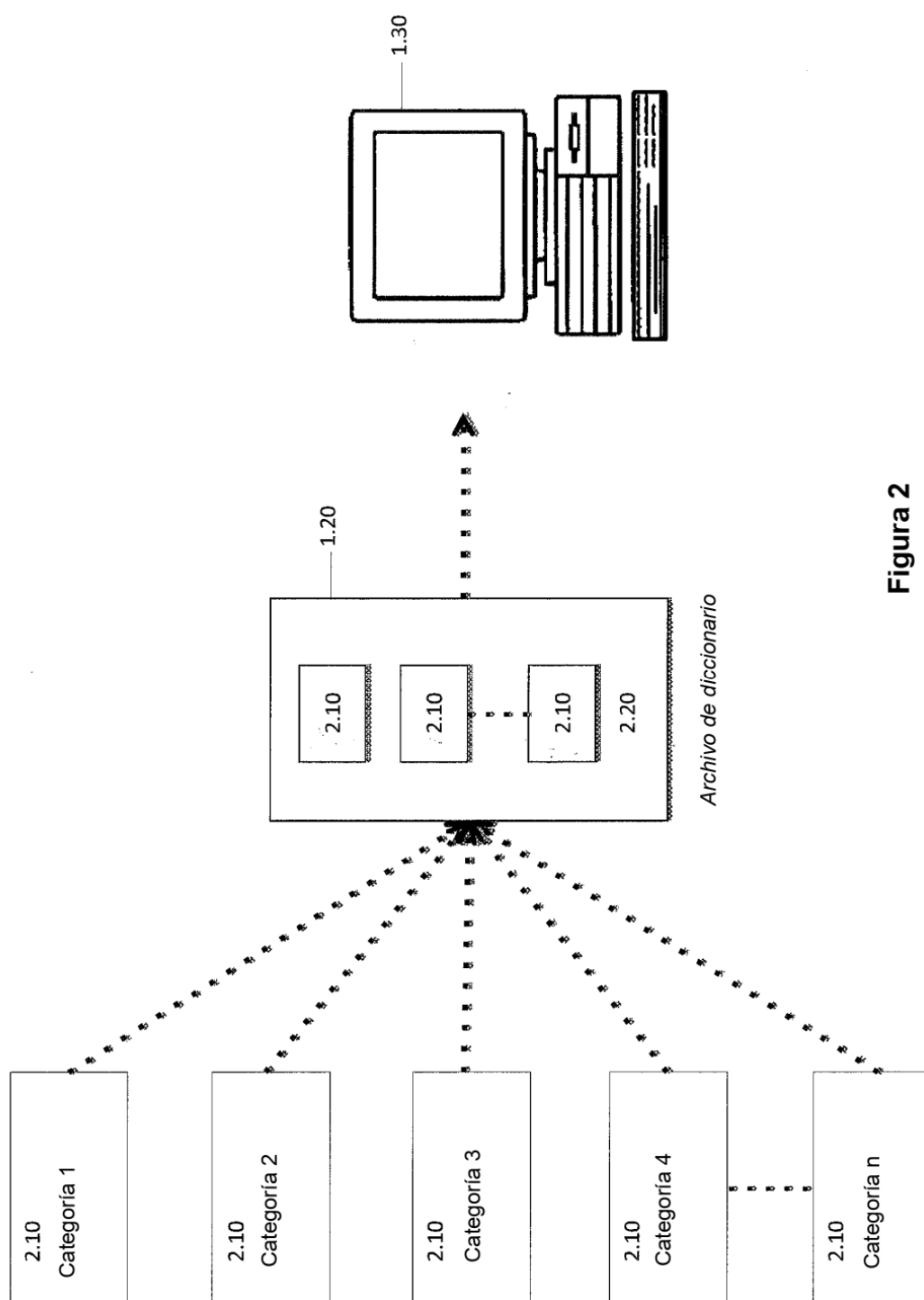


Figura 2

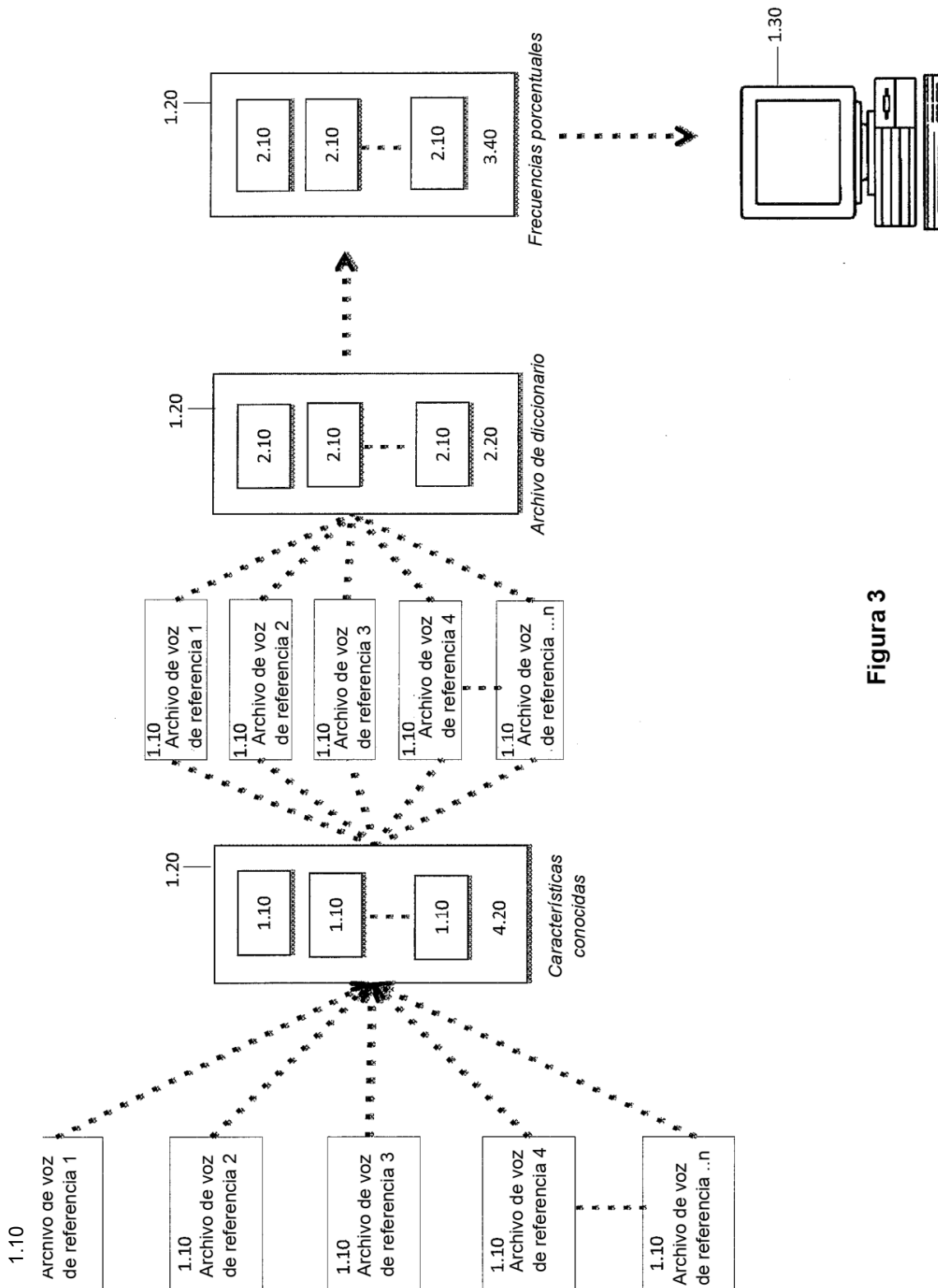


Figura 3

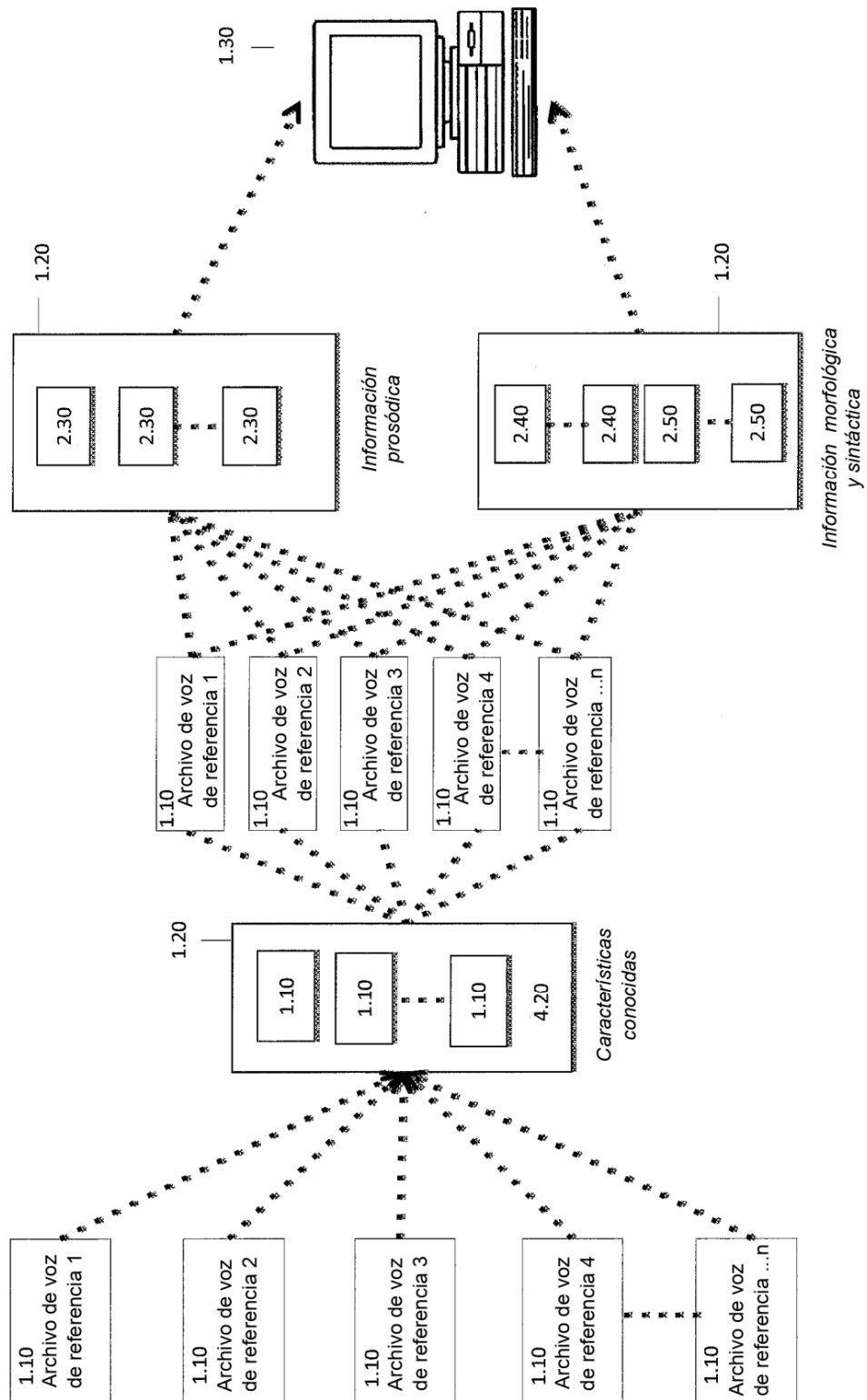


Figura 4

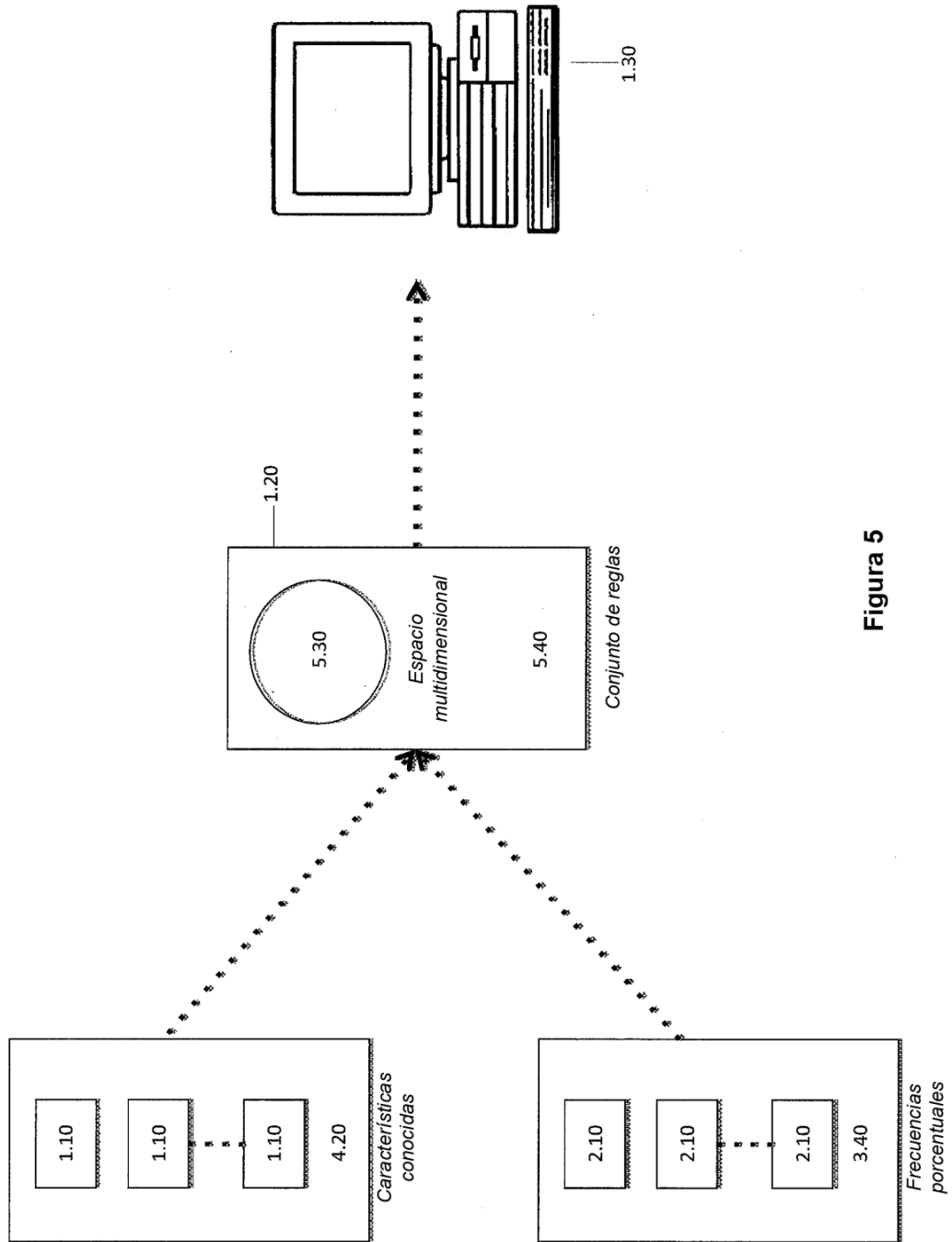


Figura 5

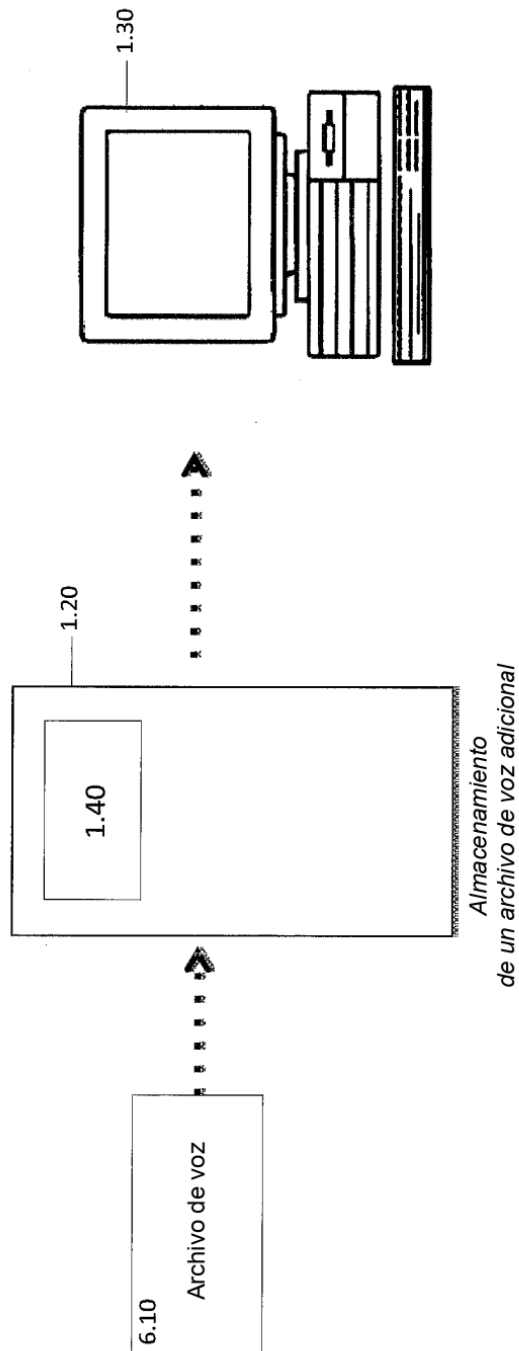


Figura 6

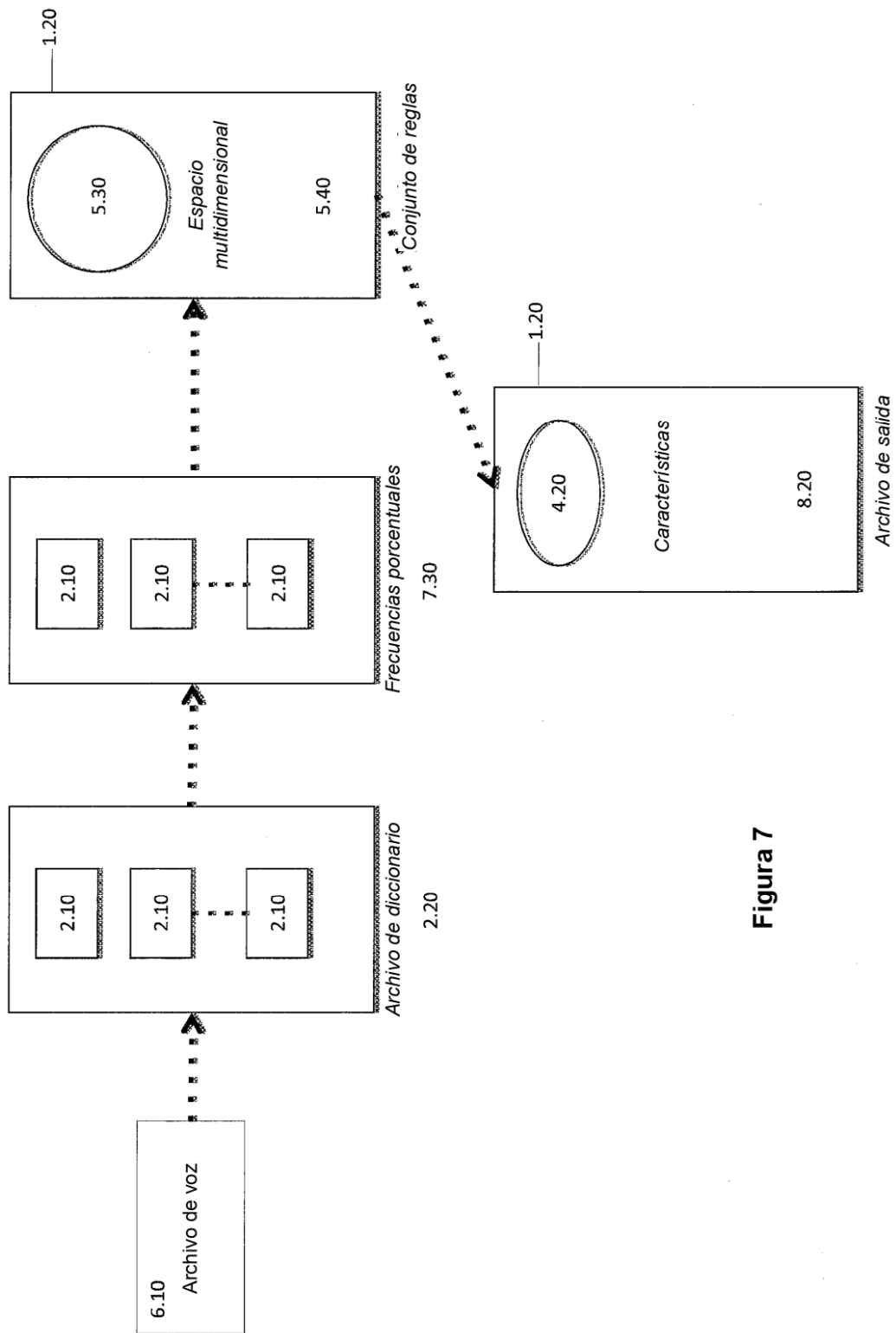


Figura 7

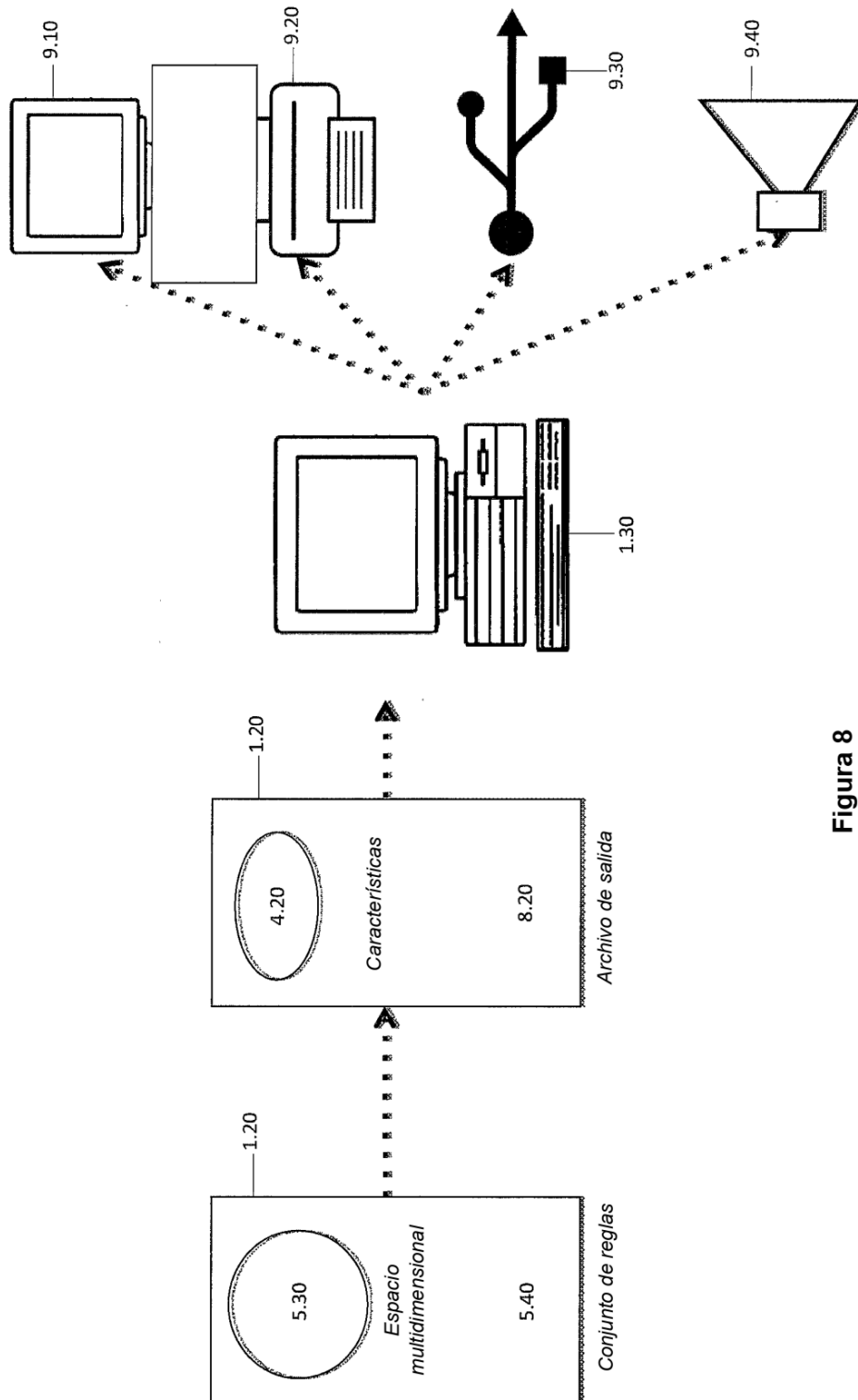


Figura 8

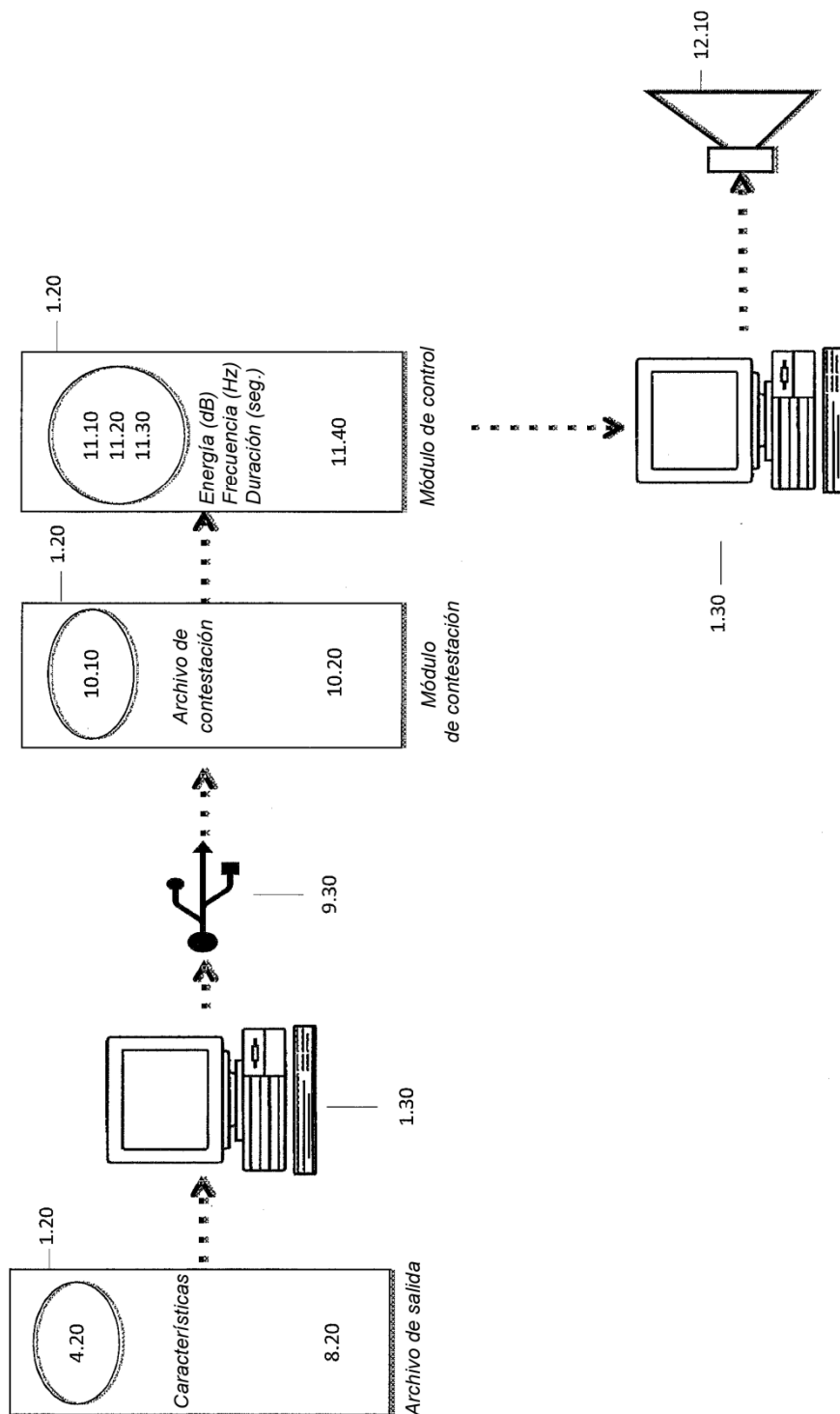


Figura 9