

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 394 726**

51 Int. Cl.:

G10L 15/18 (2006.01)

G06F 19/00 (2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **18.08.2005 E 05789851 (2)**

97 Fecha y número de publicación de la solicitud europea: **23.05.2007 EP 1787288**

54 Título: **Extracción automática de contenido semántico y generación de un documento estructurado a partir del habla**

30 Prioridad:

20.08.2004 US 923517

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

05.02.2013

73 Titular/es:

**MULTIMODAL TECHNOLOGIES, LLC (100.0%)
1710 Murray Avenue
Pittsburgh, PA 15217 , US**

72 Inventor/es:

**FRITSCH, JUERGEN;
FINKE, MICHAEL;
KOLL, DETLEF;
WOSZCZYNA, MONIKA y
YEGNANARAYANAN, GIRIJA**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 394 726 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

Extracción automática de contenido semántico y generación de un documento estructurado a partir del habla

Referencia cruzada a solicitudes relacionadas

5 Esta solicitud está relacionada con la Solicitud de patente de los Estados Unidos titulada "Documento Transcription System Training".

Antecedentes

Campo de la invención

La presente invención se refiere al reconocimiento automático del habla, y más en particular, a técnicas para transcribir automáticamente el habla.

10 **Técnica relacionada**

Es deseable en muchos contextos generar un documento escrito basado en el habla humana. En la profesión legal, por ejemplo, los transcribidos transcriben testimonios dados en procedimientos judiciales y en declaraciones para producir una transcripción escrita del testimonio. Asimismo, en la profesión médica, se producen transcripciones de diagnósticos, pronósticos, prescripciones y otras informaciones dictadas por los médicos y otros profesionales médicos. Las transcripciones en estos y otros campos necesitan típicamente ser muy precisas (medidas en términos de grado de correspondencia entre el contenido semántico (significado) del discurso original y el contenido semántico de la transcripción resultante) debido a la confianza puesta en las transcripciones resultantes y el perjuicio que podría causar una imprecisión (tal como proporcionar una prescripción de fármaco incorrecta a un paciente). Los altos grados de fiabilidad pueden, sin embargo ser difíciles de obtenerse de manera consistente por varias razones, tales como las variaciones en : (1) las características de los hablantes cuyo habla es transcrito (por ejemplo, acento, volumen, dialecto, velocidad); (2) condiciones externas (por ejemplo, ruido de fondo); (3) el transcriptor o el sistema de transcripción (por ejemplo capacidades de escucha o captura de audio imperfectas, comprensión imperfecta del lenguaje); o (4) medio de grabación/transmisión (por ejemplo, papel, cinta de audio analógica, red telefónica analógica, algoritmos de compresión aplicados en redes telefónicas digitales, y ruidos/artefactos debidos a los canales de teléfonos celulares).

25 En un primer momento la transcripción solo se realizaba por transcribidos humanos que escuchaban el habla, el discurso, bien entiendo real (por ejemplo, en persona "tomando dictado") o escuchando una grabación. Una ventaja de los transcribidos humanos es que pueden tener un conocimiento específico de un campo, tal como el conocimiento de la medicina y la terminología médica, lo cual les permite interpretar ambigüedades en el discurso y por lo tanto mejorar la precisión de la transcripción. Los transcribidos humanos, sin embargo, tienen varios inconvenientes. Por ejemplo, los transcribidos humanos producen transcripciones a una velocidad relativamente lenta y su precisión se va reduciendo a lo largo del tiempo como consecuencia del cansancio.

35 Existen varios sistemas de reconocimiento automático del habla para reconocer el habla humano generalmente y para transcribir el habla en particular. Los sistemas de reconocimiento de habla que crean transcripciones son denominados "sistemas de transcripción automatizados" o "sistemas de dictado automatizados". El software de disco listo para usar, por ejemplo, puede ser usado por los usuarios de un ordenador personal para dictar documentos en un procesador de texto como alternativa a escribir tales documentos usando un teclado.

Los sistemas de dictado automatizados intentan típicamente producir una transcripción del discurso palabra a palabra. Tal transcripción, en la cual hay una correspondencia unívoca entre las palabras en el flujo de audio hablado y las palabras en la transcripción, se denomina en el presente documento "transcripción literal". Los sistemas de dictado automatizados no son perfectos y por lo tanto pueden fallar al producir transcripciones literalmente perfectas.

45 En algunas circunstancias, sin embargo, una transcripción literal no es deseable. De hecho, los transcribidos pueden intencionalmente introducir varios cambios en la transcripción escrita. Una transcripción puede, por ejemplo, filtrar efectos espontáneos del habla (por ejemplo expresiones de pausa, vacilaciones, comienzos falsos), descartar observaciones y comentarios irrelevantes, convertir datos en un formato estándar, insertar encabezamientos u otros materiales explicativos, o cambiar la secuencia del discurso para ajustar la estructura de un informe escrito.

En el campo médico, por ejemplo, los informes hablados producidos por los médicos se transcriben a menudo en informes escritos con formatos estándar. Por ejemplo con referencia a la figura 1B, se muestra un ejemplo de un informe médico estructurado y formateado 111. El informe 111 incluye una variedad de secciones 112-138 que aparecen en una secuencia predeterminada cuando se visualiza el informe 111. En el ejemplo particular mostrado en la figura 1B, el informe incluye una sección de encabezado 112, una sección subjetiva 122, una sección objetiva 134, una sección de evaluación 136, y una sección de plan 138. Las secciones pueden incluir texto así como subsecciones. Por ejemplo, la sección de encabezado 112 incluye una sección de nombre de hospital 120 (que contiene el texto "Hospital General"), una sección de nombre de paciente 114 (que contiene el texto "Jane DOE", una sección de número de tarjeta 116 (que contiene el texto "851D"), y una sección de fecha de informe 118 (que contiene el texto (10/1/1993").

Asimismo, la sección subjetiva 122 incluye varias informaciones subjetivas acerca del paciente, incluidas tanto en el texto como en una sección de historial médico 124, una sección de medicaciones 126, una sección de alergias 128, una sección de historial familiar 130, y una sección de historial social 132. La sección objetiva 134 incluye varias informaciones objetivas. Aunque no se ilustra en la figura 1B, la información en la sección objetiva puede incluir subsecciones que contienen la información ilustrada. La sección de evaluación 136 incluye una evaluación textual de la condición del paciente, y la subsección de plan 138 incluye una descripción textual de un plan de tratamiento.

Cabe resaltar que la información puede aparecer en una forma diferente en el informe 111 de la forma en que tal información fue dictada por el médico. Por ejemplo la fecha en la sección de fecha de informe 118 puede haber sido dicha como "octubre, uno de 1993" "primero de octubre de 1993" o de otra forma. El transcriptor, sin embargo, transcribió tal discurso usando el texto "10/1/1993) en la sección de fecha de informe 118, quizás porque el hospital especificado en la sección de hospital 120 requiere que las fechas se expresen en los informes escritos con tal formato.

Asimismo, la información en la informe médico 111 puede no aparecer en la misma secuencia que en la grabación de audio original, debido a la necesidad de conformarse a un formato de informe requerido o por algún otro motivo. Por ejemplo, el médico que dicta puede haber dictado en primer lugar la sección objetiva 134, seguida de la sección subjetiva 122, y a continuación el encabezado 120. El informe escrito 111, sin embargo, contiene el encabezado 120 en primer lugar, seguido de la sección subjetiva 122 y a continuación la sección objetiva 134. Tal estructura de informe puede, por ejemplo, ser necesaria para los informes médicos en el hospital especificado en la sección de hospital 120.

El principio del informe 111 puede haber sido generado basándose en un flujo de audio hablado tal como el siguiente: "El doctor Smith uh el primero de octubre um de 1993, identidad del paciente ochentaicinco uno d um a continuación se encuentra el historial familiar del paciente que he revisado" Debería ser evidente que una transcripción literal de este discurso sería difícil de entender y no sería particularmente útil.

Cabe resaltar por ejemplo que algunas palabras, tales como " a continuación se encuentra un" no aparecen en el informe escrito 111. Asimismo, la expresión que marca pausa como "uh" no aparece en el informe escrito 111. Además, el informe escrito 111 organiza el discurso original en las secciones predefinidas 112-140 reordenando el discurso. Como lo ilustran estos ejemplos, el informe escrito 111 no es una transcripción literal del discurso de médico que dicta.

En resumen, un informe tal como el informe 111 puede ser más deseable que una transcripción literal por varios motivos (por ejemplo, porque organiza la información de tal manera que facilita la comprensión). Por lo tanto sería deseable que un sistema de transcripción automática fuese capaz de generar un informe estructurado (en lugar de una transcripción literal) basándose en el discurso sin estructurar.

Con referencia a la figura 1A, se muestra un diagrama de flujo de datos de un sistema de la técnica anterior 100 para generar un documento estructurado 110 basado en un flujo de audio hablado 102. Tal sistema produce el documento de texto estructurado 110 a partir del flujo de audio hablado 102 usando un procedimiento en dos etapas: (1) un reconocedor de habla automático 104 genera una transcripción literal 106 basada en el flujo de audio hablado 102; y (2) un procesador de lenguaje natural 108 identifica la estructura en la transcripción 106 y de este modo crea el documento estructurado 110, que tiene el mismo contenido que la transcripción 106, pero que está organizado dentro de la estructura (por ejemplo formato de informe) identificada por el procesador de lenguaje natural 108.

Por ejemplo, algunos sistemas existentes intentan generar documentos de texto estructurados ; (1) analizando el flujo de audio hablado 102 para identificar y distinguir el contenido hablado en el flujo de audio 102 de pistas estructurales explícitas o implícitas en el flujo de audio 102; (2) convirtiendo las porciones de "contenido" del flujo de audio hablado 102 en texto en bruto; y (3) usando las pistas estructurales identificadas para convertir el texto en bruto en el informe estructurado 110. Ejemplos de pistas estructurales explícitas incluyen instrucciones de formato (por ejemplo, "párrafo nuevo", "línea nueva", "siguiente punto") e identificadores de párrafos (por ejemplo "constataciones", impresiones" conclusiones"). Ejemplos de pistas estructurales implícitas incluyen pausas largas que pueden indicar límites de párrafo, indicios prosódicos que indican el fin de una enumeración, y el propio contenido hablado.

Por varios motivos descritos más en detalle en lo sucesivo, el documento estructurado 110 producido por el sistema 100 puede que no sea óptimo. Por ejemplo, el documento estructurado 110 puede contener palabras incorrectamente transcritas (por ejemplo mal reconocidas), la estructura del documento estructurado 110 puede fallar en reflejar la estructura deseada del documento, y el contenido del flujo de audio hablado 102 puede ser insertado en las subestructuras incorrectas (por ejemplo, secciones, párrafos o frases) en el documento estructurado.

Asimismo, además de o en lugar de generar el documento estructurado 110 basado en el flujo de audio hablado 102, puede ser deseable extraer el contenido semántico (tal como información acerca de medicaciones, alergias o enfermedades previas del paciente descritas en el flujo de audio 102) del flujo de audio hablado 102. Aunque tal contenido semántico puede ser útil para generar el documento estructurado 110, tal contenido puede también ser útil para otros fines, tales como popularizar una base de datos de información de paciente que se puede analizar independientemente del documento 110. Los sistemas de la técnica anterior tal como el sistema 100 mostrado en la figura 1, sin embargo, están típicamente destinados a generar el documento estructurado 110 basándose principalmente o solo en información sintáctica en el flujo de audio hablado 102. Tales sistemas no son, por lo tanto, útiles para extraer contenido semántico.

Lo que se necesita es sin embargo, técnicas mejoradas para generar documentos estructurados basados en flujos de audio hablados.

5 El documento US 2002/0123891 divulga un procedimiento de conversión de habla en texto usando una jerarquía de modelos contextuales. En el procedimiento divulgado, el modelo contextual que refleja más precisamente una o más expresiones habladas de usuario se usa para convertir el discurso en texto. El preámbulo de las reivindicaciones independientes anexas al presente documento se basa en este documento.

Sumario

10 Se divulgan técnicas para generar automáticamente documentos estructurados basados en el habla, incluyendo identificación de conceptos relevantes y su interpretación. En una realización, un generador de documento estructurado usa un procedimiento integrado para generar un documento de texto estructurado (tal como un informe médico de texto estructurado) basado en un flujo de audio hablado. El flujo de audio hablado puede ser reconocido usando un modelo lingüístico que incluye una pluralidad de submodelos dispuestos en una estructura jerárquica. Cada uno de los submodelos puede corresponder a un concepto que está previsto que aparezca en el flujo de audio hablado. Por ejemplo, los submodelos pueden corresponder a secciones de documento. Los submodelos pueden, por ejemplo, ser 15 modelos lingüísticos de n-gramas o gramáticas sin contexto. Diferentes porciones del flujo de audio hablado pueden ser reconocidas usando diferentes submodelos. El documento de texto estructurado resultante puede tener una estructura jerárquica que corresponde a la estructura jerárquica de los submodelos lingüísticos que se usaron para generar el documento de texto estructurado.

20 Por ejemplo, en un aspecto de la presente invención, se proporciona un procedimiento de acuerdo con la reivindicación independiente 1.

Otras características y ventajas de varios aspectos y realizaciones de la presente invención se pondrán de manifiesto a partir de la siguiente descripción y de las reivindicaciones.

Breve descripción de los dibujos

25 La figura 1A es un diagrama de flujo de datos de un sistema de la técnica anterior para generar un documento estructurado basado en un flujo de audio hablado;
 La figura 1B ilustra un informe médico de texto generado basado en un informe hablado;
 La figura 2 es un diagrama de flujo de un procedimiento que se ejecuta en una realización de la presente invención para generar un documento de texto estructurado basado en un documento hablado;
 30 La figura 3 es un diagrama de flujo de datos de un sistema que lleva a cabo el procedimiento de la figura 2 en una realización de la presente invención;
 La figura 4 ilustra un ejemplo de un flujo de audio hablado en una realización de la presente invención;
 La figura 5 ilustra un documento de texto estructurado de acuerdo con una realización de la presente invención;
 La figura 6 es un ejemplo de un documento presentado que es presentado basado en el documento de texto estructurado de la figura 5 de acuerdo con una realización de la presente invención
 35 La figura 7 es un diagrama de flujo de un procedimiento que es ejecutado por el generador de documento estructurado de la figura 3 en una realización de la presente invención para generar un documento de texto estructurado.
 La figura 8 es un diagrama de flujo de datos que ilustra una porción del sistema de la figura 3 en detalle relacionado con el procedimiento de la figura 7 de acuerdo con una realización de la presente invención.
 40 La figura 9 es un diagrama que ilustra correspondencias entre modelos de lenguaje, subestructuras de documento que corresponden a los modelos de lenguaje, y contenidos candidatos producidos usando los modelos lingüísticos de acuerdo con una realización de la presente invención;
 La figura 10A es un diagrama que ilustra un modelo lingüístico jerárquico de acuerdo con una realización de la presente invención;
 45 La figura 10B es un diagrama que ilustra una trayectoria a través del modelo lingüístico jerárquico de la figura 10A de acuerdo con una realización de la presente invención;
 La figura 10C es un diagrama que ilustra un modelo lingüístico jerárquico de acuerdo con una realización de la presente invención;
 50 La figura 11A es un diagrama de flujo de un procedimiento que es ejecutado por el generador de documento estructurado de la figura 3 para generar un documento de texto estructurado de acuerdo con una realización de la presente invención;
 La figura 11B es un diagrama de flujo de un procedimiento que usa un proceso integrado para seleccionar una trayectoria a través de un modelo lingüístico jerárquico y para generar un documento de texto estructurado basado en el habla de acuerdo con una realización de la presente invención;
 55 Las figuras 11C-11D son diagramas de flujo de procedimientos que son ejecutados en una realización de la presente invención para calcular una puntuación de pertinencia para un documento candidato;
 La figura 12A es un diagrama de flujo de datos que ilustra una porción del sistema de la figura 3 en detalle relacionado con el procedimiento de la figura 11A de acuerdo con una realización de la presente invención.

La figura 12B es un diagrama de flujo de datos que ilustra una realización del generador de documento estructurado de la figura 3 que lleva a cabo el procedimiento de la figura 11B en una realización de la presente invención.

5 La figura 13 es un diagrama de flujo de un procedimiento que es usado en una realización de la presente invención para generar un modelo lingüístico jerárquico para su uso en la generación de documentos de texto estructurados.

La figura 14 es un diagrama de flujo de un procedimiento que es usado en una realización de la presente invención para generar un documento de texto estructurado que usa distintas etapas de reconocimiento de habla y análisis estructural; y

10 La figura 15 es un diagrama de flujo de un sistema que aplica el procedimiento de la figura 14 de acuerdo con una realización de la presente invención.

Descripción detallada

15 Con referencia a la figura 2, se muestra un diagrama de flujo de un procedimiento 200 que es aplicado en una realización de la presente invención para generar un documento de texto estructurado basado en un documento hablado. Con referencia a la figura 3, se muestra un diagrama de flujo de datos de un sistema 300 para aplicar el procedimiento 200 de la figura 2 de acuerdo con una realización de la presente invención.

20 El sistema 300 incluye un flujo de audio hablado 302, que puede, por ejemplo, ser un flujo de audio hablado en directo o grabado de un informe médico dictado por un médico. Con referencia a la figura 4, se muestra una representación de texto de un ejemplo del flujo de audio hablado 302. En la figura 4, el texto entre los signos porcentuales representa la puntuación hablada (por ejemplo "%coma%", "%punto%", y "%dos puntos%") indicios estructurales explícitos (por ejemplo "%nuevo párrafo%") en el flujo de audio 302. Se puede observar a partir del flujo de audio 302 ilustrado en la figura 4 que una transcripción literal del flujo de audio 302 no sería particularmente útil con el fin de entender el diagnóstico, pronóstico u otra información contenida en el informe médico representado por el flujo de audio 302.

25 El sistema 300 incluye asimismo un modelo lingüístico probabilista 304. El término "modelo lingüístico probabilista" usado en el presente documento se refiere a cualquier modelo lingüístico que asigna probabilidades a las secuencias de palabras habladas. Las gramáticas libres de contexto (probabilista) y los modelos lingüísticos de n-gramas 306a-e son ambos ejemplos de "modelos lingüísticos probabilista" tal como se usa este término en el presente documento.

30 En general, una gramática libre de contexto especifica una pluralidad de formas habladas para un concepto y asocia probabilidades a cada una de las formas habladas. Una gramática de estado finito es un ejemplo de una gramática sin contexto. Por ejemplo, una gramática de estado finito para la fecha del 1 de octubre 1993, puede incluir la forma hablada de octubre primero de 1993" con una probabilidad de 0,7, la forma hablada diez noventa y tres" con una probabilidad de 0,2 y la forma hablada "primer de octubre noventa y tres" con una probabilidad del 0,1. La probabilidad asociada a cada forma hablada es una probabilidad estimada de que el concepto será hablado en esa forma hablada en un flujo de audio particular. Una gramática de estado finito, es por lo tanto, un tipo de modelo lingüístico probabilista.

35 En general, un modelo lingüístico de n-grama especifica la probabilidad de que se produzca una secuencia particular de n palabras en un flujo de audio hablado. Se considera, por ejemplo, un modelo lingüístico "unigrama", para el cual n=1. Para cada palabra en un lenguaje, un unigrama especifica la probabilidad de que se produzca la palabra en un documento hablado. Un modelo lingüístico "bigrama" (para el cual n=2) especifica probabilidades de que se produzcan pares de palabras en un documento hablado. Por ejemplo, un modelo bigrama puede especificar la probabilidad condicional de que se produzca la palabra "gato" en un documento hablado dado que la palabra anterior en el documento era "el". Asimismo, un modelo lingüístico "trigrama" especifica las probabilidades de tres palabras y así sucesivamente. Las probabilidades especificadas por modelos lingüísticos de n-grama y gramáticas de estado finito se pueden obtener formando tales documentos usando un discurso de entrenamiento y un texto de entrenamiento, como se describe más en detalle en la solicitud de patente referenciada anteriormente titulada "Document Transcription System Training".

40 El modelo lingüístico probabilista 304 incluye una pluralidad de submodelos 306a-e, cada uno de los cuales es un modelo lingüístico probabilista. Los submodelos 306a-e pueden incluir modelos lingüísticos de n-grama. Los submodelos 306a-e puede incluir modelos lingüísticos de n-grama y/o gramáticas de estado finito en una combinación. Asimismo, como se describe más en detalle en lo sucesivo, cada uno de los submodelos 306a-e puede contener submodelos adicionales, y así sucesivamente. Aunque se muestran cinco submodelos en la figura 3, el modelo lingüístico probabilista 304 puede incluir cualquier número de submodelos.

45 El objetivo del sistema 300 mostrado en la figura 3 es producir un documento de texto estructurado 310 que incluye el contenido del flujo de audio hablado 302, en el que el contenido está organizado en una estructura particular y donde los conceptos son identificados e interpretados en una forma legible por máquina. El documento de texto estructurado 310 incluye una pluralidad de subestructuras 312a-f tal como secciones, párrafos, y/o frases. Cada una de las subestructuras 312a-f puede incluir subestructuras adicionales, y así sucesivamente. Aunque se muestran seis subestructuras en la figura 3, el documento de texto estructurado 310 puede incluir cualquier número de subestructuras.

Por ejemplo, con referencia a la figura 5, se muestra un ejemplo del documento de texto estructurado 310. En el ejemplo ilustrado en la figura 5, el documento de texto estructurado 310 es un documento XML. El documento de texto estructural 310 puede, sin embargo, ser aplicado en cualquier forma. Como se muestra en la figura 5, el documento estructurado 310 incluye seis subestructuras 312a-f, cada una de las cuales puede representar una sección del documento 310.

5 Por ejemplo, el documento estructurado 310 incluye la sección de encabezado 312a que incluye metadatos acerca del documento 310, tal como un título 314 del documento 310 (“Exploración CT del tórax sin contraste”) y la fecha 316 en la cual se dictó el documento 310 (“<fecha>22-abr-2003</fecha>”). Obsérvese que el contenido en la sección de encabezado 312a se obtuvo desde el principio del flujo de audio hablado 302 (Figura 4). Asimismo, cabe resaltar que la
10 sección de encabezado 312a incluye tanto texto plano (por ejemplo, el título 314) como una subestructura (por ejemplo, la fecha 316) que representa un concepto de que se ah interpretado de una forma legible por ordenador como un triplete de valores (día-mes-año).

Representar la fecha de una forma legible por ordenador permite que la fecha sea almacenada en una base de datos y sea procesada más fácilmente que si la fecha fuese almacenada en forma de texto. Por ejemplo, si múltiples fechas en el flujo de audio 302 han sido reconocidas y almacenadas en forma legible por máquina, tales fechas pueden compararse
15 fácilmente entre sí por un ordenador. En otro ejemplo, la información estadística acerca del contenido del flujo de audio 302, tal como el tiempo medio entre las visitas médicas, pueden generarse fácilmente si las fechas están almacenadas en forma legible por ordenador. Esta ventaja de realizaciones de la presente invención se aplica generalmente no solo a fechas sino al reconocimiento de cualquier tipo de contenido semántico y el almacenamiento de tal contenido en forma legible por máquina.

20 El documento estructurado 310 incluye, además, una sección de comparación 312b, que incluye contenido que describe estudios anterior llevados a cabo en el mismo paciente que el paciente que es objeto del documento (informe) 310. Cabe destacar que el contenido en la sección de comparación 312b se obtuvo a partir de la porción del flujo de audio 302 que empieza con “comparación con” y termina con “seis de abril de dos mil uno”, pero que la sección de comparación 312b no incluye el texto “comparación con” que está en un ejemplo de un indicio de sección. El uso de tales
25 indicios para identificar el principio de una sección u otra subestructura de documento se describirá más en detalle en lo sucesivo.

En resumen, el documento estructurado 310 incluye una sección técnica 312c, que describe técnicas que han sido llevadas a cabo en los procedimientos realizados sobre el paciente; una sección de constataciones 312d, que describe las constataciones del médico; y una sección de impresión 312e, que describe las impresiones del médico acerca del
30 paciente.

Los documento XML, tal como el documento estructural ejemplar 310 ilustrados en la figura 5, no están destinados típicamente para ser visto directamente por un usuario final. En su lugar, tales documentos son representados típicamente de una manera que es más fácil de leer antes de ser presentados al usuario final. El sistema 300, por ejemplo, incluye un motor de presentación 314 que presenta el documento de texto estructurado 310 basado en una hoja
35 de estilo 316 para producir un documento presentado 318. Las técnicas para generar hojas de estilo y para presentar documentos según las hojas de estilo son bien conocidas por los expertos en la técnica.

Con referencia a la figura 6, se muestra un ejemplo del documento presentado 318. El documento presentado 318 incluye cinco secciones 602a-e, cada una de las cuales puede corresponder a una o más de las seis subestructuras 312a-f en el documento de texto estructurado 310. Más específicamente, el documento presentado 318 incluye una
40 sección de encabezado 602a, una sección de comparación 602b, una sección técnica 602c, una sección de constataciones 602d y una sección de impresiones 602e. Cabe resaltar que puede haber o no una correspondencia unívoca entre secciones en el documento presentado 318 en el documento de texto estructurado 310. Por ejemplo, cada una de las subestructuras 312a-f no necesita representar un tipo distinto de sección de documento. Si por ejemplo, dos o más subestructuras 312a-f representan el mismo tipo de sección (tal como una sección de encabezado), el motor de
45 presentación 314 puede presentar ambas subestructuras en la misma sección del documento presentado 318.

El sistema 300 incluye un generador de documentos estructurados 308, que identifica el modelo lingüístico probabilista 304 (etapa 202), y utiliza el modelo lingüístico 304 para reconocer el flujo de audio hablado 302 y de este modo producir el documento de texto estructurado 310 (etapa 204). El generador de documento estructurado 308 puede, por ejemplo, incluir un descodificador de reconocimiento de habla automático 320 que produce cada una de las subestructuras 312a-f
50 en el documento de texto estructurado 310 que usa un submodelo correspondiente de los submodelos 306a-e en el modelo lingüístico probabilista 304. Como es bien sabido por el experto en la técnica, un descodificador es un componente de un reconocedor de habla que convierte audio en texto. El descodificador 320 puede, por ejemplo, producir la subestructura 312a usando el submodelo 306a para reconocer una primera porción del flujo de audio hablado 302. Asimismo, el descodificador 320 puede producir la subestructura 312b usando el submodelo 306b para reconocer
55 una segunda porción del flujo de audio hablado 302.

Cabe resaltar que no hay necesidad de una correspondencia unívoca entre los submodelos 306a-e en el modelo lingüístico 304 y las subestructuras 312a-f en el documento estructurado 310. Por ejemplo, el descodificador de reconocimiento del habla puede usar el submodelo 306a para reconocer una primera porción del flujo de audio hablado 302 y de este modo producir las subestructuras 312a, y usar el mismo submodelo 306a para reconocer una segunda

porción del flujo de audio hablado 302 y de este modo producir la subestructura 312b. En tal caso, múltiples subestructuras en el documento de texto estructurado 310 pueden contener el contenido de una sola estructura semántica (por ejemplo, sección o párrafo).

5 El submodelo 306a puede, por ejemplo, ser un modelo lingüístico de “encabezado” que se usa para reconocer porciones del flujo de audio hablado 302 que contienen contenido en la sección de encabezado 312a; el submodelo 306b puede, por ejemplo, ser un modelo lingüístico de “comparación” que se usa para reconocer porciones del flujo de audio hablado 302 que contiene contenido en la sección de comparación 312b; y así sucesivamente. Cada modelo lingüístico de este tipo puede ser entrenado usando texto de entrenamiento de la sección correspondiente de los documentos de entrenamiento. Por ejemplo, el submodelo de encabezado 306a puede ser entrenado usando texto de las secciones de encabezado de una pluralidad de documentos de entrenamiento, y el submodelo de comparación puede ser entrenado usando texto de las secciones de comparación de la pluralidad de documentos de entrenamiento.

15 Habiéndose descrito en general características de varias realizaciones de la presente invención, las realizaciones de la presente invención se describirán ahora con más detalle. Con referencia a la figura 7, se muestra un diagrama de flujo de un procedimiento que es aplicado por el generador de documentos estructurados 308 en una realización de la presente invención para generar el documento de texto estructurado 310 (Figura 2, etapa 204). Con referencia a la figura 8, se muestra un diagrama de flujo de datos que ilustra una porción del sistema 300 en detalle pertinente para el procedimiento de la figura 7.

20 En el ejemplo ilustrado en la figura 8, el generador de documentos estructurados 308 incluye un identificador de segmentos 814 que identifica una pluralidad de segmentos S 802a-c en el flujo de audio hablado 302 (etapa 701). Los segmentos 802a-c pueden, por ejemplo, representar conceptos tales como secciones, párrafos, frases, palabras, fechas, horas, o códigos. Aunque solos se muestran tres segmentos 802a-c en la figura 8, el flujo de audio hablado 302 puede incluir cualquier número de porciones. Aunque para facilitar la explicación, todos los segmentos 802a-c son identificados en la etapa 701 de la figura 7 antes de realizar el resto del procedimiento 700, la identificación de los segmentos 802a-c puede realizarse concurrentemente con el reconocimiento del flujo de audio 302 y la generación del documento estructurado 310, como se describirá más en detalle en lo sucesivo respecto de las figuras 11B y 12B.

25 El generador de documentos estructurados 308 introduce un bucle en cada segmento S en el flujo de audio hablado 302 (etapa 702). Como se ha descrito anteriormente, el generador de documentos estructurados 308 incluye el descodificador de reconocimiento de habla 320, que puede, por ejemplo, incluir uno o más descodificadores convencionales de reconocimiento de habla que incluyen diferentes modelos lingüísticos. Además como se ha descrito anteriormente, cada uno de los submodelos 306a-e puede ser un modelo lingüístico de n-gramas, una gramática sin contexto, o una combinación de los mismos.

30 Se supone a título de ejemplo que el generador de documentos estructurados 308 está actualmente procesando el segmento 802a del flujo de audio hablado 302. El generador de documentos estructurados 308 selecciona una pluralidad 804 de submodelos 306a-e con los cuales reconocer el actual segmento S. Los submodelos 804 pueden por ejemplo, ser todos los submodelos lingüísticos 306a-e o un subconjunto de los submodelos 306a-e. El descodificador de reconocimiento de habla 320 reconoce el segmento actual S (por ejemplo, el segmento 802a) con cada uno de los submodelos seleccionados 804, produciendo de este modo una pluralidad de contenidos candidatos 808 que corresponden al segmento S (etapa 704). Dicho de otro modo, cada uno de los contenidos candidatos 808 se produce usando el descodificador de reconocimiento de habla 320 para reconocer el segmento actual S que usa un submodelo distinto de los submodelos 804. Obsérvese que cada uno de los contenidos candidatos 808 puede incluir no solo texto reconocido sino también otros tipos de contenidos tales como conceptos (por ejemplo, fechas, horas, códigos, medicaciones, alergias, signos vitales, etc.). Codificados en forma legible por máquina.

35 El generador de documentos estructurados 308 incluye un selector de contenido final 810 que selecciona uno de los contenidos candidatos 808 como contenido final 812 para el segmento S (706). El selector de contenido final 810 puede usar cualquier técnica de una variedad de técnicas que son bien conocidas por el experto en la técnica para seleccionar el resultado del reconocimiento de habla que más se acerca al habla del cual se deriva.

40 El generador de documentos estructurados 308 hace un seguimiento del submodelo que es usado para producir cada uno de los contenidos candidatos 808. Se n a título de ejemplo que los submodelos 304 incluyen todos los submodelos 306a-e, y que los contenidos candidatos 808 incluye por lo tanto cinco contenidos candidatos por segmento 802a-c (uno producido usando cada uno de los submodelos 306a-e). Por ejemplo, con referencia a la figura 9, se muestra un diagrama que ilustra correspondencias entre las subestructuras de documento 312a-f, los submodelos 306a-e, y los contenidos candidatos 808a-e. Como se ha descrito anteriormente, cada uno de los submodelos 306a-e puede asociarse a una o más subestructuras correspondientes 312a-f en el documento de texto estructurado 310. Estas correspondencias son indicadas en la figura 9 mediante las correspondencias 902a-e entre las subestructuras 312a-e y los submodelos 306a-e. El generador de documentos estructurado 308 puede mantener tal es correspondencias 902a-e en una tabla o usar otros medios.

45 Cuando el descodificador de reconocimiento de habla 320 reconoce el segmento S (por ejemplo, el segmento 802a) con cada uno de los submodelos 306a-e, produce el contenido candidato correspondiente 808a-e. Por ejemplo, el contenido candidato 808a es el texto que es producido cuando el descodificador de contenido candidato 320 reconoce el segmento

802a con el submodelo 306a, el contenido candidato 808b es el texto que es producido cuando el descodificado de reconocimiento de habla 320 reconoce el segmento 802a con el submodelo 306b, y así sucesivamente. El generador de documento estructurados 308 puede registrar la correspondencia entre los contenidos candidatos 808a-e y los submodelos correspondientes 306a-e en un conjunto de correspondencias de modelo candidato-contenido 816.

5 Por lo tanto, cuando el generador de documento estructurados 308 selecciona uno de los contenidos candidatos 808a-e como el contenido final 812 para el segmento S (etapa 706), un identificador de correspondencias finales 818 puede usar las correspondencias 816 y el contenido final seleccionado 812 para identificar el submodelo lingüístico que produjo el contenido candidato que ha sido seleccionado como el contenido final 812 (etapa 708). Por ejemplo, si el contenido candidato 808c es seleccionado como el contenido final 812, se puede observar en la figura 9 que el identificador de correspondencias finales 818 puede identificar el submodelo 306C como el submodelo que produjo contenido candidato 808c. El identificador de correspondencias finales 818 puede acumular cada submodelo identificado en el conjunto de correspondencias 820, de manera que en cualquier momento dado las correspondencias 820 identifican la secuencia de submodelos lingüísticos que se usaron para generar los contenidos finales que han sido seleccionados para su inclusión en el documento de texto estructurado 310.

15 Una vez identificado el submodelo correspondiente al contenido final 812, el generador de documentos estructurado 308 puede identificar la subestructura de documentos asociada al submodelo identificado (etapa 710). Por ejemplo, si el submodelo 306c ha sido identificado en la etapa 708, se puede ver en la figura 9 que la subestructura de documento 312c está asociada al submodelo 306c.

20 Un insertador de contenido estructurado 822 inserta el contenido final 812 en la subestructura identificada del documento de texto estructurado 310 (etapa 712). Por ejemplo, si la subestructura 312c está identificada en la etapa 710, el insertador de texto 514 inserta el contenido final 812 en la subestructura 312c.

25 El generador de documentos estructurados repite las etapas 704-712 para el resto de los segmentos 802b-c del flujo de audio hablado 302 (etapa 714), generando de este modo el contenido final 812 para cada uno de los segmentos restantes 802b-c e insertando el contenido final 812 en las subestructuras apropiadas de las subestructuras 312a-f del documento de texto 310. Al concluir el procedimiento 700, el documento de texto estructurado 310 incluye texto que corresponde al flujo de audio hablado 302, y las correspondencias finales de modelo-contenido 820 identifican la secuencia de submodelos lingüísticos que fueron usadas por el descodificador de reconocimiento de habla 320 para generar el texto en el documento de texto estructurado 310.

30 Cabe señalar que en el proceso de reconocimiento del flujo de audio hablado 302, el procedimiento 700 puede no solo generar texto que corresponde al audio hablado, sino que puede también identificar información semántica representada por el audio y almacenar tal información semántica en una forma legible por máquina. Por ejemplo, con referencia de nuevo a la figura 5, la sección de comparación 312b incluye un elemento de fecha en el cual se representa una fecha particular como un triplete que contiene valores individuales para el día ("06"), mes ("ABR"), y el año ("2001"). Otros ejemplos para conceptos semánticos en el campo médico incluyen signos vitales, medicaciones y sus dosificaciones, alergias, códigos médicos, etc. La extracción y representación de la información semántica de esta manera facilita el proceso de aplicación del procesamiento automatizado sobre tal información. Cabe señalar que la forma particular de representar la información semántica en la figura 5 es meramente un ejemplo y no constituye una limitación de la presente invención.

40 Como se recordará de la etapa 701, el procedimiento 700 mostrado en la figura 7A identifica el conjunto de segmentos 802a-c antes de identificar los submodelos a usar para reconocer los segmentos 802a-c. Cabe remarcar, sin embargo, que el generador de documentos estructurados 308 puede integrar el proceso de identificación de los segmentos 802a-c con el proceso de identificación de los submodelos a usar para reconocer los segmentos 802a-c, y con el proceso de aplicación de reconocimiento de habla sobre los segmentos 802a-c. Ejemplos de técnicas que se pueden usar para aplicar tal reconocimiento y segmentación integrada se describirán más en detalle en lo sucesivo respecto de las figuras 11B y 12B.

50 Habiendo descrito en general la operación del procedimiento ilustrado en la figura 7, se considera ahora la aplicación del procedimiento de la figura 7 al flujo de audio ejemplar 302 mostrado en la figura 4. Se supone que la primera porción del flujo de audio hablado 302 es el flujo hablado de las expresiones "Exploración CT del tórax sin contraste veintidós de abril de dos mil tres". Esta porción puede ser seleccionada en la etapa 702 y reconocida usando todos los submodelos lingüísticos 306a-e en la etapa 704 para producir una pluralidad de contenidos candidatos 808a-e. Como se ha descrito anteriormente, suponiendo que el submodelo 306a-e es un modelo lingüístico de "encabezado", que el submodelo 306b es un modelo lingüístico de "comparación", que el submodelo 306c es un modelo lingüístico de técnica, que el submodelo 306d es un modelo lingüístico de "constataciones", y el que submodelo 306e es un modelo lingüístico de "impresiones"

55 Debido al hecho que el submodelo 306a es un modelo lingüístico que ha sido entrenado para reconocer el habla en la sección de "encabezado" del documento 310 (por ejemplo, la subestructura 312a), es probable que el contenido candidato 808a producido usando el submodelo 306a coincida con las palabras de la porción de audio anteriormente mencionada en mayor medida que los otros contenidos candidatos 808b-e. Suponiendo que el contenido candidato 808a es seleccionado como el contenido final 812 para esta porción de audio, el insertador de contenido 822 insertará el

contenido final 812 producido por el submodelo 306a en la sección de encabezado 312a del documento de texto estructurado 310.

5 Se supone que la segunda porción del flujo de audio hablado es el flujo hablado de expresiones “comparación con los estudios anteriores del seis de marzo dos mil dos y el seis de abril de dos mil uno”. Esta porción se puede seleccionar en la etapa 702 y reconocer usando todos los submodelos lingüísticos 306a-e en la etapa 704 para producir una pluralidad de contenidos candidatos 808a-e. Debido al hecho que el submodelo 306b es un modelo lingüístico que ha sido entrenado para reconocer habla en la sección de “comparación” del documento 310 (por ejemplo, la subestructura 312b), es probable que el contenido candidato 808b producido usando el submodelo 306b coincida con las palabras de la porción de contenidos candidatos en mayor medida que los otros contenidos referenciados anteriormente 808a y 808c-e.

10 Suponiendo que el contenido candidato 808b es seleccionado como el contenido final 812 para esta porción de audio, el insertador de texto 514 insertará el contenido final 812 producido por el submodelo 306 en la sección de comparación 312b del documento de texto estructurado 310.

15 El resto del flujo de audio 302 ilustrado en la figura 4 puede ser reconocido e insertado en las subestructuras apropiadas de las subestructuras 312a-f en el documento de texto estructurado 310 de manera similar. Obsérvese que aunque el contenido del flujo de audio hablado 302 ilustrado en la figura 4 aparece en la misma secuencia que las secciones 312a-f en el documento de texto estructurado 310, no es una condición de la presente invención. En su lugar, el contenido puede aparecer en el flujo de audio 302 en cualquier orden. Cada uno de los segmentos 802a-c del flujo de audio 302 es reconocido por el descodificador de reconocimiento de habla 320, y el contenido final resultante 812 es insertado en la subestructura apropiada de las subestructuras 312a-f. En consecuencia, el orden del contenido de texto en las subestructuras 312a-f puede no ser el mismo que el orden del contenido en el flujo de audio hablado. Cabe señalar, sin embargo, que incluso si el orden del contenido de texto es el mismo tanto en el flujo de audio 302 que en el documento de texto estructurado 310, el motor de presentación 314 (figura 3) puede presentar el contenido de texto del documento 310 en cualquier orden deseado.

20

25 En otra realización de la presente invención, el modelo lingüístico probabilista 304 es un modelo lingüístico jerárquico. En particular, en esta realización la pluralidad de submodelos 306a-e están organizados en una jerarquía. Como se ha descrito anteriormente, los submodelos 306a-e puede incluir, además, submodelos adicionales, et así sucesivamente, de manera que la jerarquía del modelo lingüístico 304 puede incluir múltiples niveles.

30 Con referencia a la figura 10A, se muestra un diagrama que ilustra un ejemplo del modelo lingüístico 304 de forma jerárquica. El modelo lingüístico 304 incluye una pluralidad de nodos 1002, 306a-e, 1006a-e, y 1010 y 1012. Los nodos cuadrados 1002, 306b-e, y 1006e y 1012 usan gramáticas probabilistas de estado finito para modelizar conceptos muy limitados (tales como el orden de secciones de informe, indicios de sección, fechas, y horas). Los nodos elípticos 306a, 1006a-d, y 1010 usa modelos lingüísticos estadísticos (de n-gramas) para modelizar un lenguaje menos limitativo.

35 El término “concepto” tal como se usa en el presente documento incluye por ejemplo, fechas, horas, números, códigos, medicaciones, historial médico, diagnósticos, prescripciones, expresiones, enumeraciones e indicios de sección. Un concepto puede ser expresado verbalmente de muchas maneras. Cada manera de expresar verbalmente un concepto particular es denominada en el presente documento como “forma hablada” del concepto. A veces se hace una distinción entre conceptos “semánticos” y conceptos “sintácticos”. El término “concepto” tal como se usa en el presente documento incluye tanto conceptos semánticos como conceptos sintácticos, pero no se limita a ninguno de ellos y no se basa ninguna definición particular de “concepto semántico” o “concepto sintáctico” o en ninguna distinción entre ambos.

40 Se considera, por ejemplo, la fecha del 1 de octubre de 1993, que es un ejemplo de un concepto ya que este término se usa en el presente documento. Las formas habladas de este concepto incluyen las expresiones habladas “primero de octubre de mil novecientos noventa y tres”, “uno de octubre del noventa y tres”, “uno guión diez guión noventa y tres”. El texto tal como “1 de octubre de 1993) y “01/10/1993” son ejemplos de “formas habladas” de este concepto.

45 Ahora se considera la frase “John Jones tiene neumonía”. Esta frase, que es un concepto tal como se utiliza este término en el presente documento, puede ser expresado verbalmente de varias maneras, tal como las expresiones habladas, “John Jones tiene neumonía” y “paciente Jones diagnosticado con neumonía”. La frase escrita “John Jones tiene neumonía” es un ejemplo de una “forma escrita” del mismo concepto.

50 Aunque los modelos lingüísticos para conceptos de bajo nivel tal como las fechas y las horas no se muestran en la figura 10A (salvo para el submodelo 1012), el modelo lingüístico jerárquico 304 puede incluir submodelos para tales conceptos de bajo nivel. Por ejemplo, los submodelos de n-gramas 306a, 1006a-d, y 1010 pueden asignar probabilidades a secuencias de palabras que representan fechas, horas y otros conceptos de bajo nivel.

55 El modelo lingüístico 304 incluye el nodo raíz 1002, que contiene una gramática de estado finito que representa las probabilidades de ocurrencia de los subnodos 306a-e del nodo 1002. El nodo raíz 1002 puede, por ejemplo, indicar probabilidades de las secciones de encabezado, comparaciones, constataciones e impresiones del documento 310 que aparecen en ordenes particulares en el flujo de audio hablado 302.

Bajando un nivel en la jerarquía del modelo lingüístico 304, el nodo 306a es un nodo de “encabezado”, que es un modelo lingüístico de n-gramas que representa probabilidades de ocurrencia de palabras en porciones del flujo de audio hablado 302 destinado a su inclusión en la sección de encabezado 312a del documento de texto estructurado 310.

5 El nodo 306b contiene una gramática de estado finito de “comparación” que representa probabilidades de ocurrencia de varias formas habladas alternativas de indicios para la sección de comparación 312b del documento de texto. La gramática de estado finito en el nodo de comparación 306 puede por ejemplo, incluir indicios tales como “comparación con”, “comparación para”, “antes es”, y estudios anteriores son”. La gramática de estado finito puede incluir una probabilidad para cada uno de estos indicios. Tales probabilidades pueden, por ejemplo estar basadas en frecuencias de uso observadas de los indicios en un conjunto de habla de entrenamiento para el mismo hablante o en el mismo campo que el flujo de audio hablado 302. Tales frecuencias pueden obtenerse, por ejemplo, usando las técnicas divulgadas en la solicitud de patente mencionada anteriormente titulada “Document Transcription System Training”.

10 El nodo de comparación 306e incluye un subnodo 1006a de “contenido de comparación, que es un modelo lingüístico de n-gramas que representa probabilidades de ocurrencia de palabras en porciones del flujo de audio hablado 302 destinado para su inclusión en el cuerpo de la sección de comparación 312b del documento de texto 310. El nodo de contenido de comparación 1006a tiene un nodo de fecha 1012 como un niño. Como se ha descrito más en detalle en lo sucesivo, el nodo de fecha 1012 es una gramática de estado finito que representa probabilidades de la fecha sea expresada verbalmente de varias maneras.

15 Los nodos 306c y 306d pueden entenderse de manera similar. El nodo 306c contiene una gramática de estado finito de “técnica” que representa probabilidades de ocurrencia de varias formas habladas alternativas de indicios para la sección técnica 312c del documento de texto 310. El nodo técnico 306c incluye un subnodo de contenido técnico” 1006b, que es un modelo de lenguaje de n-gramas que representa probabilidades de ocurrencia de palabras en porciones del flujo de audio hablado 302 destinado para su inclusión en el cuerpo de la sección técnica 312c del documento de texto 310. 20 Asimismo, el nodo 306d contiene una gramática de estado finito de “constataciones” que representa probabilidades de ocurrencia de varias formas habladas alternativas de indicios para la sección de constataciones 312d del documento de texto 310. L nodo de constataciones 306d incluye un submodo de “contenido de constataciones” 1006c, que es un modelo lingüístico de n-gramas que representan probabilidades de ocurrencia de palabras en porciones del flujo de audio hablado 302 destinado para su inclusión en el cuerpo de la sección de constatación 312d del documento de texto 310.

25 El nodo de impresiones 306 es similar a los nodos 306b-d-, porque incluye una gramática de estado finito 1006 que incluye un modelo lingüístico de n-gramas para reconocer indicios de sección y un submodo 1006d que incluye un modelo lingüístico de n-gramas para reconocer el contenido de secciones. Además, sin embargo, el nodo de impresiones 306e incluye un submodo adicional 1006e que a su vez incluye un submodo 1010. Esto indica que el contenido de la sección de impresiones puede ser reconocida usando bien el modelo lingüístico en el nodo de contenido de impresiones 1006d o el nodo “enum” 1006e, gobernado por el modelo lingüístico basado en la gramática de estado finito que corresponde al nodo de impresiones 306e. El nodo “enum” 1006e contiene una gramática de estado finito que indica probabilidad asociadas a diferente maneras de expresar verbalmente indicios de enumeración (tal como “número uno”, “número dos”, “primero”, segundo”, “tercero”, y así sucesivamente). El nodo de contenido de impresiones 1010 puede 30 incluir el mismo modelo lingüístico como el nodo de contenido de impresiones 1006d.

35 Habiéndose descrito la estructura jerárquica del modelo lingüístico 304 en una realización de la presente invención, ahora se describirán ejemplos de técnicas que se pueden usar para generar el documento estructurado 310 usando el modelo lingüístico 304. Con referencia a la figura 11A, se muestra un diagrama de flujo de un procedimiento que se aplica por el generador de documento estructurada 308 en una realización de la presente invención 308 en una realización de la presente invención para generar el documento de texto estructurado 310 (figura 2, etapa 204). Con 40 referencia a la figura 12A, se muestra un diagrama de flujo de datos que ilustra una porción del sistema 300 en detalle pertinente para el procedimiento de la figura 11A.

EL generador de documento estructurada 308 incluye un selector de trayectoria 1202 que identifica una trayectoria 1204 a través del modelo lingüístico jerárquico 304 (etapa 1102). La trayectoria 1204 es una secuencia ordenada de nodos en el modelo lingüístico jerárquico 304. Los nodos pueden ser tiempos múltiples atravesados en la trayectoria 1204. 45 Ejemplos de técnicas para generar la trayectoria 1204 se describirán más en detalle en lo sucesivo respecto de las figuras 11B y 12B.

Con referencia en la figura 10B, se ilustra un ejemplo de la trayectoria 1204. La trayectoria 1204 incluye los puntos 1020a-j, que especifican una secuencia en la cual para atravesar nodos en el modelo lingüístico 304. Los puntos 1020a-j se denominan “puntos” en lugar de “nodos” para distinguirlos de los nodos 1002, 306a-e, 1006a-e y 1010 en el modelo 50 lingüístico 304.

En el ejemplo ilustrado en la figura 10B, la trayectoria 1204 atraviesa los siguientes nodos del modelo lingüístico 304 en la secuencia: (1) nodo raíz 1002 (punto 1020a); (2) nodo de contenido de encabezado 306a (punto 1020b); (3) nodo de comparación 306b (punto 1020c); (4) nodo de contenido de comparación 1006a (punto 1020d); (5) nodo técnico 306c (punto 1020e); (6) nodo de contenido técnico 1006b (punto 1020f); (7) nodo de constataciones 306d (punto 1020g); (8) 55 nodo de contenido de constataciones 1006c (punto 1020h); (9) nodo de impresiones 306e (punto 1020i); y (10) nodo de contenido de impresiones 1006d (punto 1020j).

Como se puede observar con referencia a la figura 4, reconocer el flujo de audio hablado 302 usando los submodelos lingüísticos que se encuentran a lo largo de la trayectoria 1204 ilustrada en la figura 10B dará como resultado el reconocimiento óptico del habla, ya que el habla en el flujo de audio 302 se produce en la misma secuencia que los

submodelos lingüísticos en la trayectoria 1204 ilustrada en la figura 10B. Por ejemplo, el flujo de audio hablado 302 empieza con el habla que es mejor reconocido por el modelo lingüístico de contenido de encabezado 306a (Exploración CT de tórax sin contraste veintidós de abril dos mil tres”), seguido de habla que es mejor reconocida por el modelo lingüístico de comparación 306b (“comparación con”), seguido por el habla que es mejor reconocida por el modelo lingüístico de contenido de comparación 1006a (antes de los estudios del seis de marzo de dos mil dos y el seis de abril de dos mil uno), y así sucesivamente.

Habiendo identificado la trayectoria 1204, el generador de documento estructurado 308 reconoce el flujo de audio hablado 302 usando los modelos lingüísticos atravesados por la trayectoria 1204 para producir el documento de texto estructurado 310 (etapa 1104). Como se describe más en detalle en lo sucesivo respecto de las figuras 11B y 12B, El reconocimiento de habla y la generación del documento de texto estructurado de la etapa 1104 se pueden integrar con la identificación de trayectoria de la etapa 1102, en lugar de realizarlos por separado..

Más específicamente, el generador de documentos estructurados 308 puede incluir un enumerador de nodos 1206 que repite cada uno de los nodos de modelo lingüístico N 1208 atravesado por la trayectoria seleccionada 1204 (etapa 1106). Para cada nodo N de este tipo, el decodificador de reconocimiento de habla 320 puede reconocer la porción del flujo de audio 302 que corresponde al modelo lingüístico en el nodo N para producir el texto estructurado T correspondiente (etapa 1108). El generador de documentos estructurado 308 puede insertar el texto T 1210 en la estructura del documento textual estructurado 310 que corresponde al nodo N 1208 del modelo lingüístico 304 (etapa 1110).

Por ejemplo, cuando el nodo N es el nodo de comparación 306b (figura 10A), el nodo de comparación 306b se puede usar para reconocer el texto “comparación con” en el flujo de audio hablado 302 (figura 4). Debido a que el nodo de comparación 306b corresponde a una subestructura de documento (por ejemplo, la sección de comparación 312b) en lugar del contenido, el resultado del reconocimiento de habla realizado en la etapa 1108 en este caso puede ser una subestructura de documento, a saber una sección de “comparación” vacía. Tal sección puede insertarse en el documento estructurado 310 en la etapa 1110, por ejemplo, en forma de etiquetas de coincidencia “<comparación>” y “/comparación>”.

Cuando el nodo N es el nodo de contenido de comparación 1006a (figura 10A), el nodo de contenido de comparación 1006a se puede usar para reconocer el texto “antes de estudios del veintiséis de marzo dos mil dos y 6 de abril de dos mil uno” en el flujo de audio hablado 302 (figura 4). Produciendo de este modo el texto estructurado “estudios anteriores de <fecha>26-MAR-2002</fecha> y <fecha>26-ABR-2001</fecha>”, como se muestra en la figura 5. Este texto estructurado se puede entonces insertar en la sección de comparación 312b en la etapa 1110 (por ejemplo, entre las etiquetas de “<comparación>” y “>/comparación>”, como se muestra en la figura 5).

El generador de documento estructurado 308 repite las etapas 1108-1110 para el resto de nodos N atravesado por la trayectoria 1204 (etapa 1112), insertando de este modo una pluralidad de textos estructurados 1210 en el documento de texto estructurado 310. El resultado final del procedimiento ilustrado en la figura 11A es la creación del documento de texto estructurado 310, que contiene texto que tiene una estructura que corresponde a la estructura de la trayectoria 1204 a través del modelo lingüístico 304. Por ejemplo, se puede observar en la figura 10B que la estructura de la trayectoria ilustrada atraviesa los nodos de modelo lingüístico que corresponden a las secciones de encabezado, comparación, técnica, constataciones e impresiones en secuencia. El documento de texto estructurado resultante 310 (como se ilustra, por ejemplo en la figura 5) incluye de manera similar las secuencias de encabezado, comparación, técnica, constataciones e impresiones en secuencia. El documento de texto estructurado 310 tiene por lo tanto la misma estructura que la trayectoria del modelo lingüístico 1204 que se usó para crear el documento de texto estructurado 310.

Se ha establecido anteriormente que el generador de documentos estructurados 308 inserta texto estructurado reconocido 1210 en las subestructuras apropiadas del documento de texto estructurado 310 (figura 11A, etapa 1110). Como se muestra en la figura 5, el documento de texto estructurado 310 se puede aplicar como un documento XML u otro documento que soporta estructuras anidadas. En tal caso, es necesario insertar cada uno de los textos estructurados reconocidos 1210 en el interior de la subestructura apropiada para que el documento de texto estructurado final 310 tenga una estructura que corresponde con la estructura de la trayectoria 1204. El experto en la técnica entenderá la manera de usar las correspondencias finales de contenido de modelo 820 (figura 8) para usar la trayectoria 1204 para atravesar la estructura del modelo lingüístico 304 y por lo tanto para crear tal documento estructurado.

El sistema ilustrado en la figura 12A incluye un selector de trayectoria 1202, que selecciona una trayectoria 1204 a través del modelo lingüístico 304. El procedimiento ilustrado en la figura 11A usa entonces la trayectoria seleccionada 1204 para generar el documento de texto estructurado 310. Dicho de otro modo, en la figura 11A y 12A, las etapas de selección de trayectoria y de creación de documento estructurado se realizan por separado. Esto no es, sin embargo, una limitación de la presente invención.

En su lugar, con referencia a la figura 11B, se muestra un diagrama de flujo de un procedimiento 1150 que integra las etapas de selección de trayectoria y generación de documentos estructurados. Con referencia a la figura 12B, se muestra una realización del generador de documentos estructurados 308 que aplica el procedimiento 1150 de la figura 11B en una realización de la presente invención. En general, el procedimiento 1150 de la figura 11B busca posibles trayectorias a través de la jerarquía del modelo lingüístico 304 (figura 10A), empezando en el nodo raíz 1002 y expandiéndose hacia fuera. Cualquiera de las diversas técnicas, incluyendo las técnicas bien conocidas por el experto en

la técnica, puede ser usada para buscar a través de la jerarquía de modelos lingüísticos. Puesto que el procedimiento 1150 identifica trayectorias parciales a través de la jerarquía de modelo lingüístico, el procedimiento 1150 usa el descodificador de reconocimiento de habla 320 para reconocer porciones cada vez mayores del flujo de audio hablado 302 usando los modelos lingüísticos que se encuentran a lo largo de las trayectorias parciales creando de este modo documentos parciales estructurados candidatos. El procedimiento 1150 asigna puntuaciones a cada uno de los documentos estructurados candidatos parciales. La puntuación de pertinencia para cada documento estructurado candidato es una medida de lo bien que se ha llevado a cabo la trayectoria que produjo el documento estructurado candidato. El procedimiento 1150 expande las trayectorias parciales siguiendo de este modo buscando a través de la jerarquía de modelo lenguaje, hasta que todo el flujo de audio hablado 302 haya sido reconocido. El generador de documentos estructurados 308 selecciona el documento estructurado candidato que tiene la mayor puntuación de pertinencia como documento de texto estructurado final 310.

Más específicamente, el procedimiento 1150 inicia una o más trayectorias candidatas 1224 a través del modelo lingüístico 304 (etapa 1152). Por ejemplo, las trayectorias candidatas 1224 pueden iniciarse para contener una única trayectoria que consiste en el nodo raíz 1002. El término "trama" se refiere en el presente documento a un periodo de tiempo corto, tal como 10 milisegundos. El procedimiento 1150 inicia un apuntador de flujo de audio para apunta a la primera trama en el flujo de audio 302 (etapa 1153). Por ejemplo, en la realización ilustrada en la figura 12B, el generador de documentos estructurados 308 contiene un enumerador de flujo de audio 1240 que proporciona una porción 1242 del flujo de audio 302 al descodificador de reconocimiento de habla 320. Al iniciarse el procedimiento 1150, la porción 1242 puede contener solo la primera trama del flujo de audio 302.

El descodificador de reconocimiento de habla 320 reconoce la porción actual 1242 del flujo de audio 302 usando los submodelos lingüísticos en la o las trayectorias candidatas 1224 para generar uno o más documentos parciales estructurados candidatos 1232 (etapa 1154). Cabe señalar que los documentos 1232 son solo documentos parciales 1232 porque se han generado basados en solo una porción del flujo de audio 302. Cuando se aplica la etapa 1154 en primer lugar, el descodificador de reconocimiento de habla 320 puede reconocer simplemente la primera trama del flujo de audio 302 usando el modelo lingüístico en el nodo raíz 1002 del modelo lingüístico 304.

Cabe señalar que las técnicas divulgadas anteriormente respecto de la figura 11A y la figura 12A pueden ser usados por el descodificador de reconocimiento de habla 320 para generar los documentos parciales estructurados candidatos 1232 usando las trayectorias candidatas 1224. Más específicamente, el descodificador de habla 320 puede aplicar el procedimiento ilustrado en la figura 11A a la porción de flujo de audio 1242 usando cada una de las trayectorias candidatas 1224 como la trayectoria identificada en la etapa 1102 (figura 11A).

Volviendo a las figuras 11B y 12B, un evaluador de pertinencia 1234 genera puntuaciones de pertinencia 1236 para cada uno de los documentos parciales estructurados candidatos 1232 (etapa 1156). Las puntuaciones de pertinencia 1236 son mediciones de lo bien que los documentos parciales estructurados candidatos 1232 representan la porción correspondiente del flujo de audio 302. En general, la puntuación de pertinencia para un documento candidato individual puede ser generada: (1) generando puntuaciones de pertinencia para cada uno de los nodos en la trayectoria correspondiente de las trayectorias candidatas 1224; y (2) usando una función de síntesis para sintetizar la puntuación de pertinencia de nodo individual generada en la etapa (1) en una puntuación de pertinencia global para el documento estructurado candidato. Ejemplos de técnicas que se pueden usar para generar las puntuaciones de pertinencia candidatas 1236 se describirán más en detalle en lo sucesivo respecto de la figura 11C.

Si el generador de documentos estructurados 308 intentase buscar todas las trayectorias posibles a través de la jerarquía del modelo lingüístico 304, los recursos informáticos necesarios para evaluar cada trayectoria posible podrían ser prohibitivos desde el punto de vista económico y del tiempo. Debido al crecimiento exponencial en el número de trayectorias posibles. Por lo tanto, en la realización ilustrada en la figura 12B, un podador de trayectoria 1230 usa las puntuaciones de pertinencia candidatas 1236 para eliminar las trayectorias mal ajustadas de las trayectorias candidatas 1224, produciendo de este modo un conjunto de trayectorias podadas 1222 (etapa 1158).

Si la totalidad del flujo de audio 302 ha sido reconocida (etapa 1160), un selector de documento final 1238 selecciona, entre los documentos parciales estructurados candidatos 1232, el documento estructurado candidato que tiene la mayor puntuación de pertinencia, y proporciona el documento seleccionado como el documento de texto estructurado final 310 (etapa 1164). Si la totalidad del flujo de audio 302 no ha sido reorganizado, un extensor de trayectoria extiende las trayectorias podadas 1222 dentro del modelo lingüístico 304 para producir un nuevo conjunto de trayectorias candidatas 1224 (etapa 1162). Si por ejemplo, las trayectorias candidatas 1222 consiste en una trayectoria individual que contiene el nodo raíz 1002, el extensor de trayectoria 1220 puede extender esta trayectoria en un nodo hacia abajo en la jerarquía ilustrada en la figura 10A para producir una pluralidad de trayectorias candidatas que se extienden desde el nodo raíz 1002, tal como una trayectoria desde el nodo raíz 1002 al nodo de contenido de encabezado 306a, una trayectoria desde el nodo raíz 1002 al nodo de comparación 306b, una trayectoria desde el nodo raíz 1002 al nodo técnico 306c, y así sucesivamente. Varias técnicas de extensión de las trayectorias 1224 para llevar a cabo búsquedas de jerarquía en profundidad, amplitud u otros tipos de búsquedas jerárquicas son bien conocidas por el experto en la técnica.

El enumerador de flujo de audio 1240 extiende la porción 1242 del flujo de audio 302 para incluir la siguiente trama en el flujo de audio 302 (etapa 1163). Las etapas 1154-1160 se repiten entonces usando las nuevas trayectorias candidatas

1224 para reconocer la porción 1242 del flujo de audio 302. De este modo la totalidad del flujo de audio 302 puede ser reconocida usando submodelos apropiados en el modelo lingüístico 304.

5 Como se ha descrito anteriormente respecto de las figuras 11B y 12B, las puntuaciones de pertinencia 1236 pueden ser generadas para cada uno de los documentos parciales estructurados candidatos 1232 producidos por el generador de documentos estructurados 308 mientras se evalúan las trayectorias candidatas 1224 a través del modelo lingüístico 304. Ahora se describirán ejemplos de técnicas para generar puntuaciones de pertinencia, bien para los documentos parciales estructurados candidatos 1232 ilustrados en la figura 12B o bien para documentos estructurado más en general.

10 Por ejemplo, con referencia a la figura 10A, cabe señalar que el nodo de contenido de comparación 1006a tiene un nodo de fecha 1012 como un nodo descendiente. Se supone que el texto "Exploración CT del tórax sin contraste veintidós de abril de dos mil tres" ha sido reconocido como texto que corresponde al nodo de contenido de comparación 1006a. Cabe señalar que el nodo de contenido de comparación 1006a se usó para reconocer el texto "Exploración CT del tórax sin contraste" y que el nodo de fecha 1012, que es un nodo descendiente del nodo de contenido de comparación 1006a, se uso para generar el texto "veintidós de abril dos mil tres". La puntuación de pertinencia para este texto puede, por lo tanto, calcularse usando el nodo de contenido de comparación 1006a para calcular una primera puntuación de pertinencia para el texto "Exploración CT del tórax sin contraste" Seguido de cualquier fecha, calculando una segunda puntuación de pertinencia para el texto "veintidós de abril dos mil tres" basado en el nodo de fecha 1012, y multiplicando la primera y segunda puntuaciones de pertinencia.

20 Con referencia a la figura 11C, se muestra un diagrama de flujo de un procedimiento que es aplicado en una realización de la presente invención para calcular una puntuación de pertinencia para un documento candidato, y que puede por lo tanto usarse para aplicar la etapa 1156 del procedimiento 1150 ilustrado en la figura 11B. Se inicia una puntuación de pertinencia S en un valor de uno para el documento estructurado candidato que se está evaluando (etapa 1172). El procedimiento asigna un apuntado de nodo actual N para apuntar al nodo raíz en la trayectoria candidata que corresponde al documento candidato (etapa 1174).

25 El procedimiento requiere una función denominada Pertinencia () con los valores N y S (etapa 1176) y restituye el resultado como la puntuación de pertinencia para el documento candidato (etapa 1178). Como se describirá ahora más en detalle, la función Pertinencia () genera la puntuación de pertinencia S usando una factorización jerárquica atravesando la trayectoria candidata que corresponde al documento candidato.

30 Con referencia a la figura 11D, se muestra un diagrama de flujo de la función Pertinencia () 1180 según una realización de la presente invención. La función 1180 identifica la probabilidad $P(W|N)$ de que el texto W que corresponde al nodo actual N ha sido reconocido por el modelo lingüístico asociado a ese nodo, y multiplica la probabilidad por el valor actual de S para producir un nuevo valor para S (etapa 1184).

35 Si el nodo N no tiene nodo descendente (etapa 1186), el valor de S es restituido (etapa 1194). Si el nodo N tiene nodo descendente, entonces se requiere la función Pertinencia () 1180 de manera recursiva encada uno de los nodos descendentes, con los resultados que se están multiplicando por el valor de S para producir nuevos valores de S (etapas 1188-1192). El valor resultando de S es restituido (etapa 1194).

Al concluir el procedimiento ilustrado en la figura 11C, el valor de S representa una puntuación de pertinencia para la totalidad del documento estructurado candidato, y el valor de S es restituido, por ejemplo, para su uso en el procedimiento 1150 ilustrado en la figura 11B (etapa 1194).

40 Por ejemplo, volviendo de nuevo al texto "Exploración CT del tórax sin contraste veintidós de abril dos mil tres". La puntuación de pertinencia (probabilidad) de este texto se puede obtener identificando la probabilidad del texto "Exploración CT de tórax sin contraste <FECHA>", donde <FECHA> indica cualquier fecha, multiplicada por la probabilidad condicional del texto "veintidós de abril dos mil tres" que se produce dado que el texto representa una fecha.

45 Más en general, el efecto del procedimiento ilustrado en la figura 11C es incorporar jerárquicamente las probabilidades de secuencias de palabras según la jerarquía del modelo lingüístico 304, permitiendo las estimaciones de probabilidad individuales asociadas a cada nodo de modelo lingüístico de combinarse sin dificultad con las estimaciones de probabilidad asociadas con otros nodos. Este marco probabilista permite que el sistema modelice y use modelos lingüístico estadísticos con gramáticas de estado finito probabilista integradas y modelos lingüísticos estadísticos integrados.

50 Como se ha descrito anteriormente, los nodos en el modelo lingüístico 304 representa submodelos lingüístico que especifican las probabilidades de ocurrencia de secuencias de palabras en el flujo de audio hablado 302. En la discusión anterior, se ha supuesto que las probabilidades ya se han asignado en tales modelos lingüísticos. Ahora se divulgarán ejemplos de técnicas para asignar probabilidades a los submodelos lingüísticos (tal como modelos lingüísticos de n-gramas y gramáticas sin contexto) en el modelo lingüístico 304.

55 Con referencia a la figura 13, se muestra un diagrama de flujo de un procedimiento 1300 que se usa en una realización de la presente invención para generar el modelo lingüístico 304. Se selecciona una pluralidad de nodos para su uso en el modelo lingüístico (etapa 1302). Los nodos pueden, por ejemplo, ser seleccionados por un transcriptor u otra persona experto en el ámbito pertinente. Los nodos pueden ser seleccionados en un intento de capturar todos los tipos de

conceptos que se pueden dar en el flujo de audio hablado 302. Por ejemplo, en el ámbito médico, se puede seleccionar nodos (tales como los mostrados en la figura 10A) que representan las secciones de un informe médico y los conector (tales como fechas, horas, medicaciones, alergias, signos vitales y códigos médicos) que se esperan que haya en un informe médico.

5 Un concepto y cualquier tipo de modelo lingüístico pueden ser asignados a cada uno de los nodos seleccionados en la etapa 1302 (etapas 1304-1306). Por ejemplo, el nodo 306b (figura 10A) puede ser asignado al concepto "indicio de sección de comparación" y ser asignado al tipo de modelo lingüístico "gramática de estado finito". De manera similar, el nodo 1006a puede ser asignado al concepto "contenido de comparación" y el tipo de modelo lingüístico "modelo lingüístico de n-gramas".

10 Los nodos seleccionados en la etapa 1302 pueden disponerse en una estructura jerárquica (etapa 1308). Por ejemplo, los nodos 1002, 306a-e, 1006a-e y 1010 pueden disponerse en la estructura jerárquica ilustrada en la figura 10A para representar y aplicar dependencias estructurales entre los nodos.

Cada uno de los nodos seleccionados en la etapa 1302 puede entonces entrenarse usando texto que representa un concepto correspondiente (etapa 1310). Por ejemplo, un conjunto de documentos de entrenamiento puede ser identificado. El conjunto de documentos de entrenamiento puede, por ejemplo, ser un conjunto de informes médicos existentes u otros documentos en el mismo ámbito como el flujo de audio hablado 302. Los documentos de entrenamiento pueden ser marcados manualmente para indicar la existencia y la ubicación de estructuras en el documento, tal como secciones, subsecciones, fechas, horas, códigos, y otros conceptos. Tal marcación puede, por ejemplo, ser realizada automáticamente sobre documentos formateados, o manualmente por transcritores u otros

15 persona cualificadas en el ámbito pertinente. Ejemplos de técnicas para entrenar los nodos seleccionados en la etapa 1302 se describen en la solicitud de patente referenciada anteriormente titulada "Document Transcription System Training"

Técnicas de entrenamiento de modelos lingüísticos convencionales pueden ser usadas en la etapa 1310 para entrenar los modelos lingüísticos de concepto específico para cada uno de los conceptos que se marcan en los documentos de

25 entrenamiento. Por ejemplo, el texto de todas las secciones de "encabezado" marcadas en los documentos de entrenamiento se puede usar para entrenar el nodo de modelo lingüístico 306a que representa la sección de encabezado. De esta manera se pueden entrenar los modelos lingüísticos para cada uno de los nodos 1002, 306a-e, 1006a-e, y 1010 en el modelo lingüístico 304 ilustrado en la figura 10A. El resultado del procedimiento 1300 ilustrado en la figura 13 es un modelo lingüístico jerárquico que tiene posibilidades de entrenamiento, que se pueden usar para

30 generar el documento de texto estructurado 310 de la manera descrita anteriormente. Este modelo lingüístico jerárquico se puede usar entonces, por ejemplo, para resegmentar repetidamente el texto de entrenamiento, como por ejemplo usando las técnicas divulgadas anteriormente e junto con las figuras 11B y 12B. El texto de entrenamiento resegmentado se puede usar para retener el modelo lingüístico jerárquico. Este proceso de resegmentación y reentrenamiento se puede aplicar repetidamente para mejorar repetidamente la calidad del modelo lingüístico.

35 En los ejemplos descritos anteriormente, el generador de documentos estructurados 308 reconoce el flujo de audio hablado 302 y genera el documento de texto estructurado 310 usando un proceso integrado, generando una transcripción intermedia no estructurada. Tales técnicas, sin embargo, están divulgadas simplemente a título de ejemplo y no constituyen limitaciones de la presente invención.

40 Con referencia a la figura 14 se muestra un diagrama de flujo de un procedimiento 1400 que se usa en otra realización de la presente invención para generar el documento de texto estructurado 310 usando distintas etapas de reconocimiento de habla y análisis estructural. Con referencia a la figura 15, se muestra un diagrama de flujo de datos de un sistema 1500 que lleva a cabo el procedimiento 1400 de la figura 14 según una realización de la presente invención.

El decodificador de reconocimiento de habla 320 reconoce el flujo de audio hablado 302 usando un modelo lingüístico 1506 para producir una transcripción 1502 del flujo de audio hablado 302. Cabe señalar que el modelo lingüístico 1506

45 puede ser un modelo lingüístico convencional que es distinto del modelo lingüístico 304. Más específicamente, el modelo lingüístico 1506 puede ser un modelo lingüístico monolítico convencional. El modelo lingüístico 1506 puede, por ejemplo, ser generado usando el mismo cuerpo de entrenamiento que el que se usa para entrenar el modelo lingüístico 304. Mientras las porciones del cuerpo de entrenamiento pueden ser usadas para entrenar el modelo lingüístico 304, el cuerpo entero puede ser usado para entrenar el modelo lingüístico 1506. El decodificador de reconocimiento de habla

50 320, puede, de este modo, usar técnicas de reconocimiento de habla convencionales para reconocer el flujo de audio hablado 302 usando el modelo lingüístico 1506 y producir de este modo la transcripción 1502.

Cabe señalar que la transcripción 1502 puede ser una transcripción "plana" 1502 del flujo de audio hablado 302, en lugar de un documento estructurado como en los ejemplos anteriores divulgados anteriormente. La transcripción 1502 puede, por ejemplo, incluir una secuencia de texto plano que se parece al texto ilustrado en la figura 4 (que ilustra el flujo de

55 audio hablado 302 en forma de texto).

El sistema 1500 incluye también un analizador estructural 1504, que usa el modelo lingüístico jerárquico 304 para analizar la transcripción 1502 y producir de este modo el documento de texto estructurado 310 (etapa 1404). El analizador estructural 1504 puede usar las técnicas divulgadas anteriormente respecto de las figuras 11C y 12B para: (1)

producir múltiples documentos estructurados candidatos que tienen el mismo contenido que la transcripción 1502 pero con estructuras que corresponden a diferentes trayectorias a través del modelo lingüístico 304; (2) generar una puntuación de pertinencia para cada uno de los documentos estructurados candidatos; y (3) seleccionar el documento estructurado candidato con la mayor puntuación de pertinencia como el documento de texto estructurado final. Contrariamente a las técnicas divulgadas anteriormente respecto de las figuras 11C y 12B, sin embargo, la etapa 1404 puede ser aplicada sin llevar a cabo el reconocimiento de habla para generar cada uno de los documentos estructurados candidatos. En su lugar, una vez producida la transcripción 1502 usando el descodificador de reconocimiento de habla 320, los documentos estructurados candidatos pueden ser generados basados en la transcripción 1502 sin llevar a cabo un reconocimiento de habla adicional.

Asimismo, el analizador estructural 1504 no necesita usar todo el modelo lingüístico 304 para producir el documento de texto estructurado 310. En su lugar, el analizador estructural 1504 puede usar un modelo lingüístico “esquelético” a escala reducida, tal como el modelo lingüístico 1030 ilustrado en la figura 10C. Cabe señalar que el modelo lingüístico ejemplar 1030 mostrado en la figura 10C es el mismo que el modelo lingüístico 304 mostrado en la figura 10A, salvo que en el modelo lingüístico esquelético 1030 los nodos de modelo lingüístico de contenido 306a, 1006a-d y 1010 han sido sustituidos por modelos lingüísticos aceptados universalmente 1032a-f, también denominados como modelo lingüístico “No importa”. Los modelos lingüísticos 1032a-f aceptarán cualquier texto que les sea proporcionado como entrada. El modelo de lenguaje de indicios de encabezado 306b-e en el modelo lingüístico esquelético 1030 permite que el analizador estructural 1504 analice la transcripción 1502 en las subestructuras correctas en el documento estructurado 310. El uso de los modelos lingüísticos aceptados universalmente 1032a-f, permite sin embargo que el analizador estructural 1504 lleve a cabo tal análisis estructural sin incurrir en el gasto (típicamente considerable) de los modelos lingüísticos de contenido de entrenamiento, tal como los modelos 306a, 1006a-d y 1010 mostrados en la figura 10A.

Cabe señalar que el modelo lingüístico esquelético 1030 puede seguir también incluir modelos lingüísticos, tales como el modelo lingüístico de fechas 1012, que corresponde a conceptos de bajo nivel. Como consecuencia, el modelo lingüístico esquelético 1030 se puede usar para generar el documento estructurado 310 de la transcripción 1502 sin incurrir en la sobrecarga de modelos lingüísticos de contenido de entrenamiento, mientras conserva la capacidad de analizar conceptos de nivel inferior en el documento estructurado 310.

Entre las ventajas de la invención se encuentra una o más de las siguientes. Las técnicas divulgadas en el presente documento sustituyen el modelo lingüístico global tradicional con una combinación de modelos lingüísticos locales especializados que están mejor adaptados a la sección de un documento que un modelo lingüístico genérico individual. Tal modelo lingüístico tiene varias ventajas.

Por ejemplo, el uso de un modelo lingüístico que contiene submodelos, cada uno de los cuales corresponde a un concepto particular, es ventajoso porque permite que el modelo lingüístico más apropiado sea utilizado para reconocer el habla que corresponde a cada concepto. Dicho de otro modo, si cada uno de los submodelos corresponde a un concepto diferente, entonces cada uno de los submodelos puede ser usado para aplicar un reconocimiento de habla en el habla que representa el concepto correspondiente. Dado que las características del habla pueden variar de un concepto a otro, el uso de tales modelos lingüísticos específicos de concepto puede producir mejores resultados de reconocimiento que los que se producirían usando un modelo lingüístico monolítico para todos los conceptos.

Aunque los submodelos de un modelo lingüístico pueden corresponder a secciones de un documento, esto no es una limitación de la presente invención. En su lugar, cada submodelo en el modelo lingüístico puede corresponder a cualquier concepto, tal como una sección, párrafo, frase, fecha, hora o código ICD9. En consecuencia, los submodelos en el modelo lingüístico pueden coincidir con conceptos particulares con un grado de precisión superior al que sería posible si solo se empleasen modelos lingüísticos específicos de sección. El uso de tales modelos lingüísticos específicos de concepto para una amplia variedad de conceptos puede mejorar, además, la precisión de reconocimiento de habla.

Asimismo, los modelos lingüísticos jerárquicos designados de acuerdo con las realizaciones de la presente invención pueden tener estructuras jerárquicas multinivel, con el efecto de anidar submodelos los unos en el interior de los otros. Como consecuencia; los submodelos en el modelo lingüístico se puede aplicar a porciones del flujo de audio hablado 302 con varios niveles de granularidad, siendo aplicado el modelo lingüístico más apropiado a cada nivel de granularidad. Por ejemplo, un modelo lingüístico de “sección de encabezado” se puede aplicar generalmente al habla en el interior de la sección de encabezado de un documento, mientras que un modelo lingüístico de “fecha” se puede aplicar específicamente al habla que representa fechas en la sección de encabezado. Esta capacidad para anidar modelos lingüísticos y aplicar modelos lingüísticos anidados a diferentes porciones de habla puede mejorar, además, la precisión de reconocimiento permitiendo que el modelo lingüístico más apropiado sea aplicado a cada porción de un flujo de audio hablado.

Otra ventaja del uso de un modelo lingüístico que incluye una pluralidad de submodelos es que las técnicas divulgadas en el presente documento pueden usar tal modelo lingüístico para generar un documento de texto estructurado a partir de un flujo de audio hablado usando un único proceso integrado, en lugar del proceso en dos etapas de la técnica anterior 100 ilustrado en la figura 1A en el que la etapa de reconocimiento de habla va seguida de una etapa de procesamiento del lenguaje natural. En el proceso en dos etapas 100 ilustrado en la figura 1A las etapas llevadas a cabo por el reconocedor de habla 104 y el procesador de lenguaje natural 108 están completamente desacoplados. Debido al reconocedor de habla automático 104 y al procesador de lenguaje natural 108 funcionan independientemente el uno del

otro, el resultado 106 del reconocedor de habla automático 104 es una transcripción literal del contenido hablado en el flujo de audio 102. La transcripción literal 106 contiene de este modo el texto correspondiente a todas las palabras habladas en el flujo de audio 102, sean estas palabras relevantes o no relevantes para el documento de texto estructurado deseado final. Tales palabras pueden incluir, por ejemplo, dudas, palabras extrañas o repeticiones, así como pistas estructurales o palabras relacionadas con la tarea. Además, el procesador lingüístico natural 108 se basa en la detección y transcripción exitosas de ciertas palabras clave y/o expresiones clave, tales como pistas estructurales. Si estas palabras/expresiones clave son mal reconocidas por el reconocedor de habla automático 104, la identificación de entidades estructurales por el procesador lingüístico natural 108 puede verse afectado negativamente. Por el contrario, en el procedimiento 200 ilustrado en la figura 2, el reconocimiento de habla y el procesamiento del lenguaje natural están integrados, permitiendo de este modo que el modelo lingüístico influya tanto en el reconocimiento de palabras en el flujo de audio 4302 como en la generación de estructura en el documento de texto estructurado 310, mejorando de este modo la calidad global del documento estructurado 310.

Además de generar el documento estructurado 310, las técnicas divulgadas en el presente documento también se pueden usar para extraer e interpretar contenido semántico del flujo de audio 302. Por ejemplo, el modelo lingüístico de fechas 1012 (figuras 10A-10B) se puede usar para identificar porciones del flujo de audio 302 que representan fechas, y almacenar representaciones de tales fechas en forma legible por ordenador. Por ejemplo, las técnicas divulgadas en el presente documento se pueden usar para identificar la expresión hablada "primero de octubre mil novecientos noventa y tres" como una fecha y almacenar la fecha de una forma legible por ordenador, tal como "mes=10, día=1, año=1998). Almacenar tales conceptos de una forma legible por ordenador permite que el contenido de tales conceptos sea procesado fácilmente por un ordenador, por ejemplo seleccionando secciones de documento por fecha o identificando medicaciones prescritas antes de una fecha dada. Además, las técnicas divulgadas en el presente documento permite que el usuario defina diferentes porciones (por ejemplo, secciones) del documento y elegir qué conceptos se han de extraer en cada sección. Las técnicas divulgadas en el presente documento facilitan de este modo el reconocimiento y el procesamiento del contenido semántico en los flujos de audio hablados. Tales técnicas se pueden aplicar en lugar de o además del almacenamiento de información extraída en un documento estructurado.

Ámbitos tales como los ámbitos médicos y legales, en los cuales hay grandes cuerpos de flujos de audio grabados preexistentes para su uso como texto de entrenamiento, pueden ser particularmente ventajosos en las técnicas divulgadas en el presente texto. Tal texto de entrenamiento se puede usar para entrenar el modelo lingüístico 304 usando las técnicas divulgadas anteriormente respecto de la figura 13. Dado que pueden ser necesarios documentos en tales ámbitos para tener estructuras bien definidas, y dado que tales estructuras se pueden identificar fácilmente en los documentos existentes, puede ser relativamente fácil (aunque consumiendo tiempo) para identificar correctamente las porciones de tales documentos específicos de concepto para su uso en el entrenamiento de cada uno de los nodos de modelo lingüístico de concepto específico en el modelo lingüístico 304. Como consecuencia, cada uno de los nodos de modelo lingüístico puede ser bien entrenado para reconocer el concepto correspondiente, aumentando de este modo la precisión de reconocimiento y aumentando la capacidad del sistema para generar documentos con la estructura requerida.

Asimismo, las técnicas divulgadas en el presente documento se pueden aplicar en tales ámbitos sin requerir ningún cambio en el proceso existente por el cual el audio es grabado y transcrito. En el ámbito médico, por ejemplo, los médicos pueden seguir dictando informes médicos a su manera habitual. Las técnicas divulgadas en el presente documento se pueden usar para generar documentos con la estructura deseada sin tener en cuenta la manera de dictar el flujo de audio hablado. Técnicas alternativas que requieren cambios en el flujo de trabajo, tales como técnicas que requieren que los hablantes se registren (leyendo texto de entrenamiento), que requieren que los hablantes modifiquen su manera de hablar (por ejemplo diciendo siempre los conceptos usando formas habladas predeterminadas), o requieren que las transcripciones sean generadas en un formato particular, pueden tener un coste prohibitivo para su aplicación en ámbitos tales como los campos médico y legal. Tales cambios pueden, de hecho, ser incoherentes con las necesidades institucionales o legales relacionadas con la estructura del informe (tal como los requisitos de informe de seguro). Las técnicas divulgadas en el presente documento, por el contrario, permiten que el flujo de audio 302 sea generado de cualquier manera y tenga cualquier forma.

Asimismo, los submodelos individuales 306a-e en el modelo lingüístico 304 se puede actualizar fácilmente sin afectar al resto del modelo lingüístico. Por ejemplo, el submodelo de contenido de encabezado 306a-e se puede sustituir por un submodelo de contenido de encabezado diferente que se representa diferentemente por la forma en que se dictó la cabecera del documento. 'La estructura modular del modelo lingüístico 304' permite que tal modificación/sustitución de submodelos sea llevado a cabo sin la necesidad de modificar ninguna parte del modelo lingüístico 304. Como consecuencia, las partes del modelo lingüístico 304 se pueden actualizar fácilmente para reflejar diferentes acuerdos de dictado de documentos.

Asimismo, el documento de texto estructurado 310 que es producido por varias realizaciones de la presente invención se puede usar para entrenar un modelo lingüístico. Por ejemplo, las técnicas de entrenamiento descritas en la solicitud de patente referenciada anteriormente titulada "Documento Transcripción System Training" puede usar el documento de texto estructurado 310 para volver a entrenar y de este modo mejorar el modelo lingüístico 304. El modelo lingüístico vuelto a entrenar 304 se puede usar entonces para producir posteriores documentos de texto estructurados, los cuales pueden a su vez ser usado para volver a entrenar el modelo lingüístico 304. Este proceso iterativo se puede emplear para mejorar la calidad de los documentos estructurados que se producen a lo largo del tiempo.

Cabe entender que aunque la invención se ha descrito anteriormente en términos de realizaciones particulares, las realizaciones anteriores están provistas solo a título ilustrativo, y no limitan o definen el alcance de la invención. Otras varias realizaciones, que incluyen pero no se limitan a lo siguiente, se encuentran también dentro del alcance de las reivindicaciones. Por ejemplo, los elementos y componentes descritos en el presente documento se pueden dividir, además, en componentes adicionales o unirse juntos para formar menos componentes para llevar a cabo las mismas funciones.

El flujo de audio hablado 302 puede ser cualquier flujo de audio, tal como un flujo de audio en directo recibido directamente o indirectamente (tal como en una conexión telefónica o por IP) o un flujo de audio grabado en cualquier medio y en cualquier formato. En el reconocimiento de habla distribuido (DSR), un cliente lleva a cabo el preprocesamiento en un flujo de audio para producir un flujo de audio procesado que es transmitido a un servidor, que lleva a cabo el reconocimiento de habla sobre el flujo de audio procesado. El flujo de audio 302 puede, por ejemplo, ser un flujo de audio procesado producido por un cliente DSR.

Aunque en los ejemplos anteriores se describe cada nodo en el modelo lingüístico 304 como conteniendo un modelo lingüístico que corresponde a un concepto particular, no es un requisito de la presente invención. Por ejemplo, un nodo puede incluir un modelo lingüístico que resulta de la interpolación de un modelo lingüístico de concepto específico asociado al nodo con uno o más de: (1) modelos lingüísticos globales de fondo con otros nodos, o (2) modelos lingüísticos específicos de concepto asociados a otros nodos.

En los ejemplos anteriores, se puede hacer una distinción entre “gramáticas” y “texto”. Cabe apreciar que el texto puede ser representado como una gramática, en la que es una única forma hablada que tiene una sola probabilidad. Por lo tanto, los documentos que se describen en el presente documento como se incluye tanto en el texto como en las gramáticas se pueden aplicar solo usando gramáticas si se desea. Además, una gramática de estado único es simplemente un tipo de gramática sin contexto, que es un tipo de modelo lingüístico que permite que múltiples formas habladas alternativas de un concepto sea aplicado más generalmente a cualquier otro tipo de gramática. Asimismo, aunque la descripción anterior puede referirse a gramáticas de estado finito y modelos lingüísticos de n-gramas, hay simplemente ejemplos de tipos de modelos lingüísticos que se pueden usar junto con las realizaciones de la presente invención. Las realizaciones de la presente invención no están limitadas a su uso junto con cualquier tipo o tipos particulares de modelo o modelos lingüísticos.

La invención no se limita a cualquiera de los campos descritos (tal como informes médicos y legales), sino que se aplica generalmente a cualquier tipo de documentos estructurados.

Las técnicas descritas anteriormente pueden aplicarse, por ejemplo, en hardware, software, microprogramas o cualquier combinación de los mismos. Las técnicas descritas anteriormente se puede aplicar en uno o más programas informáticos que se ejecutan en un ordenador programable incluyendo un procesador, un medio de almacenamiento legible por el procesador (incluyendo, por ejemplo, elementos de almacenamiento y/o memoria volátil y no volátil), al menos un dispositivo de entrada, y al menos un dispositivo de salida. El código de programa se puede aplicar a la entrada introducida usando el dispositivo de entrada para llevar a cabo las funciones descritas y generar la salida. La salida puede ser provista a uno o más dispositivos de salida.

Cada programa informático dentro de las reivindicaciones siguientes puede ser aplicado en cualquier lenguaje de programación, tal como lenguaje ensamblador, lenguaje máquina, un lenguaje de programación de procedimiento de alto nivel, o un lenguaje de programación orientado al objeto. El lenguaje de programación puede, por ejemplo, ser un lenguaje de programación compilado o interpretado.

Cada programa informático de este tipo se puede aplicar en un producto de programa informático materializa de manera tangible en un dispositivo de almacenamiento legible por máquina para su ejecución por un procesador informático. Las etapas del procedimiento de la invención pueden ser llevadas a cabo por un procesador informático que ejecuta un programa materializado de manera tangible en un medio legible por ordenador para aplicar las funciones de la invención que funcionan en la entrada y generan la salida. Los procesadores apropiados incluyen, a título de ejemplo, tanto microprocesadores de uso general como de uso especial. Generalmente el procesador recibe instrucciones y datos de una memoria de solo lectura y/o una memoria de acceso aleatorio. Los dispositivos de almacenamiento apropiados para materializar de manera tangible instrucciones de programa informático incluyen, por ejemplo, todas formas de memoria no volátil, tal como dispositivos de memoria semiconductora, incluyendo dispositivos EPROM, EEPROM, y de memoria ultrarrápida; discos magnéticos tales como los discos duros internos y discos removibles; discos magneto-ópticos; y CD-ROM. Cualquiera de los anteriores puede complementarse con, o incorporarse a, ASIC especialmente diseñadas (Circuitos integrados de aplicación específica o FPGA (Matrices de puertas programables por campos). Un ordenador puede en general recibir también programas y datos de un medio de almacenamiento tal como un disco interno (no mostrado) o un disco removible). Estos elementos se encontrarán también en un ordenador de sobremesa convencional o de estación de trabajo así como otros ordenadores apropiados para ejecutar programas informáticos que aplican los procedimientos descritos en el presente documento, que se pueden usar junto con cualquier motor de impresión digital o motor de marcado, monitor de visualización, u otro dispositivo de salida de trama capaz de producir píxeles de color o escala de grises sobre papel, película, pantalla de visualización, u otro medio de salida.

REIVINDICACIONES

1.- Un procedimiento que comprende las etapas:

- (A) identificar un modelo lingüístico probabilista que incluye una pluralidad de modelos lingüísticos probabilistas asociada a una pluralidad de subestructuras de un documentos; y
- (B) utilizar un descodificador de reconocimiento de habla para aplicar el modelo lingüístico probabilista a un flujo de audio hablado para producir un documento que incluye contenido organizado en la pluralidad de subestructuras, en el cual el contenido en cada una de la pluralidad de subestructuras es producido reconociendo el habla usando la subestructura, en el cual la pluralidad de modelos lingüísticos probabilistas están organizados en una jerarquía, y en el cual la etapa (B) comprende las etapas de:

(B)(1) identificar una trayectoria a través de la jerarquía, que comprende las etapas de:

(B)(1)(a) identificar una pluralidad de trayectoria a través de la jerarquía

(B)(1)(b) para cada una de la pluralidad de trayectorias P, producir un documento estructurado candidato para el flujo de audio hablado usando el descodificador de reconocimiento de habla para reconocer el flujo de audio hablado usando los modelos lingüísticos en la trayectoria P;

B(1)© aplicar una medición a la pluralidad de documentos estructurados candidatos producidos en la etapa (B)(1)(b) para producir una pluralidad de puntuaciones de pertinencia para la pluralidad de documentos estructurados candidatos; y

(B)(1)(d) seleccionar la trayectoria que produce los documentos estructurados candidatos que tienen la mayor puntuación pertinente;

(B)(2) generar el documento que tiene una estructura que corresponde a la trayectoria identificada en la etapa (B)(1).

2.- El procedimiento de la reivindicación 1, en el cual la pluralidad de modelos lingüísticos probabilistas incluye al menos un modelo lingüístico de n-gramas.

3.- El procedimiento de la reivindicación 1, en el cual la pluralidad de modelos lingüísticos probabilistas incluye al menos un modelo lingüístico de estado finito.

4.- El procedimiento de la reivindicación 1, en el cual la pluralidad de subestructura incluye una subestructura que representa un concepto semántico.

5.- El procedimiento de la reivindicación 4, en el cual el concepto semántico comprende una medicación.

6.- El procedimiento de la reivindicación 1, que comprende, además, una etapa de:

(C) presentar el documento para producir una representación que indica la estructura del documento.

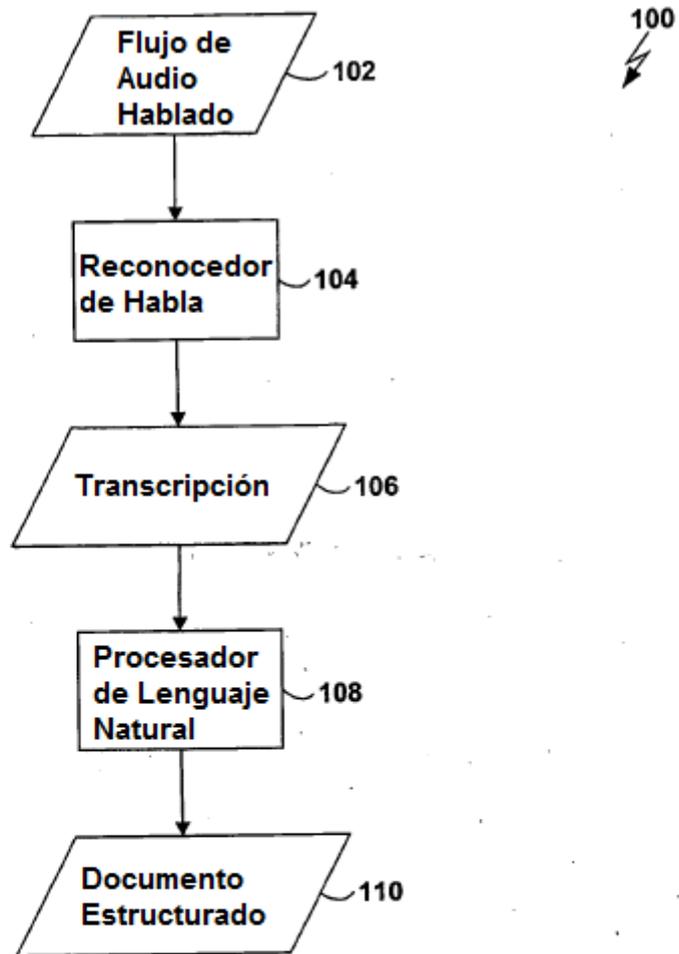


FIG. 1A (Técnica Anterior)

111 ↘

112	HOSPITAL GENERAL	118	Fecha: 10/01/1993
114	Nombre de Paciente: JANE DOE		
116	Nº de Tarjeta: 851D		
122	<p>S: La Sra XXX acude hoy quejándose de un escozor y secreción acuosa en el ojo derecho. La paciente relata que esto viene sucediendo desde hace varios meses</p>		
	124	H MP:	Significante para histerectomía, BSO...
	126	Med:	Estratest 1 mg. c.d.
	128	Alergias:	NKDA.
	130	H Fam:	Revisado.
	132	H Soc:	Revisado.
134	O:	WT: 140 T: 87.6 P: 84	BP: 116/64. GENERAL - La paciente es una
136	A:	WDWN, mujer blanca sentada sin sufrimiento agudo	
		I. Copnjuntivitis, ojo derecho. Se sospecha aletgia, no se puede excluir glaucoma	
138	P:	<p>1.- Se receta gotas oculares Nombre de la Marca: 2 gotas en el ojo derecho tres veces al día 2.- Se programa una cita de la paciente con el Dr. Dyer, su optometrista para la evaluación de su ojo derecho y excluir la posibilidad de glaucoma 3.- La paciente volverá si los síntomas no mejoran, en caso contrario volverá el día de la cita programada</p>	

FIG. 1B

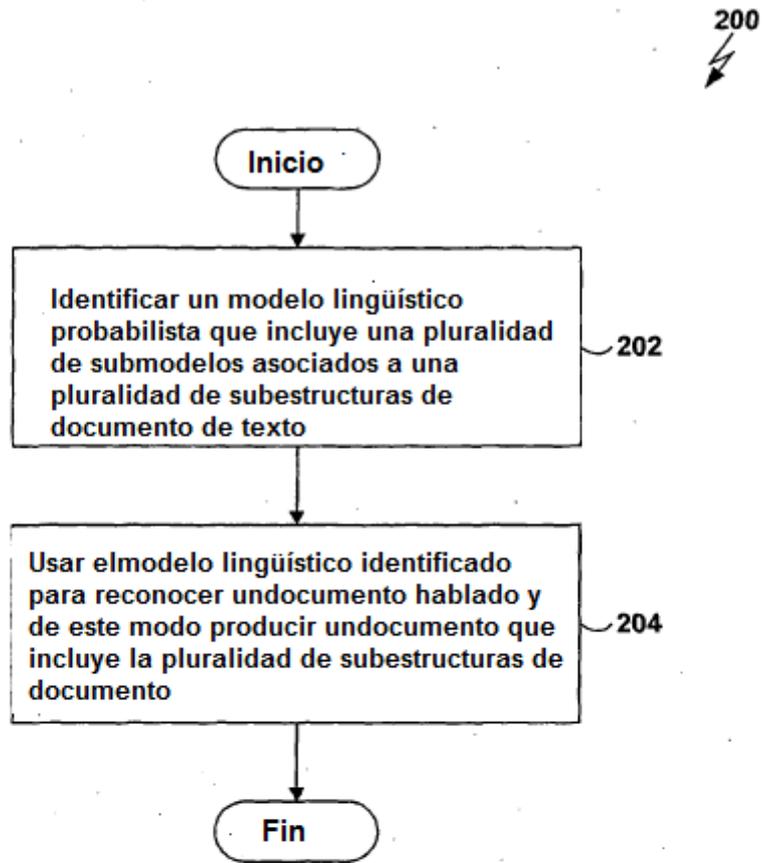


FIG. 2

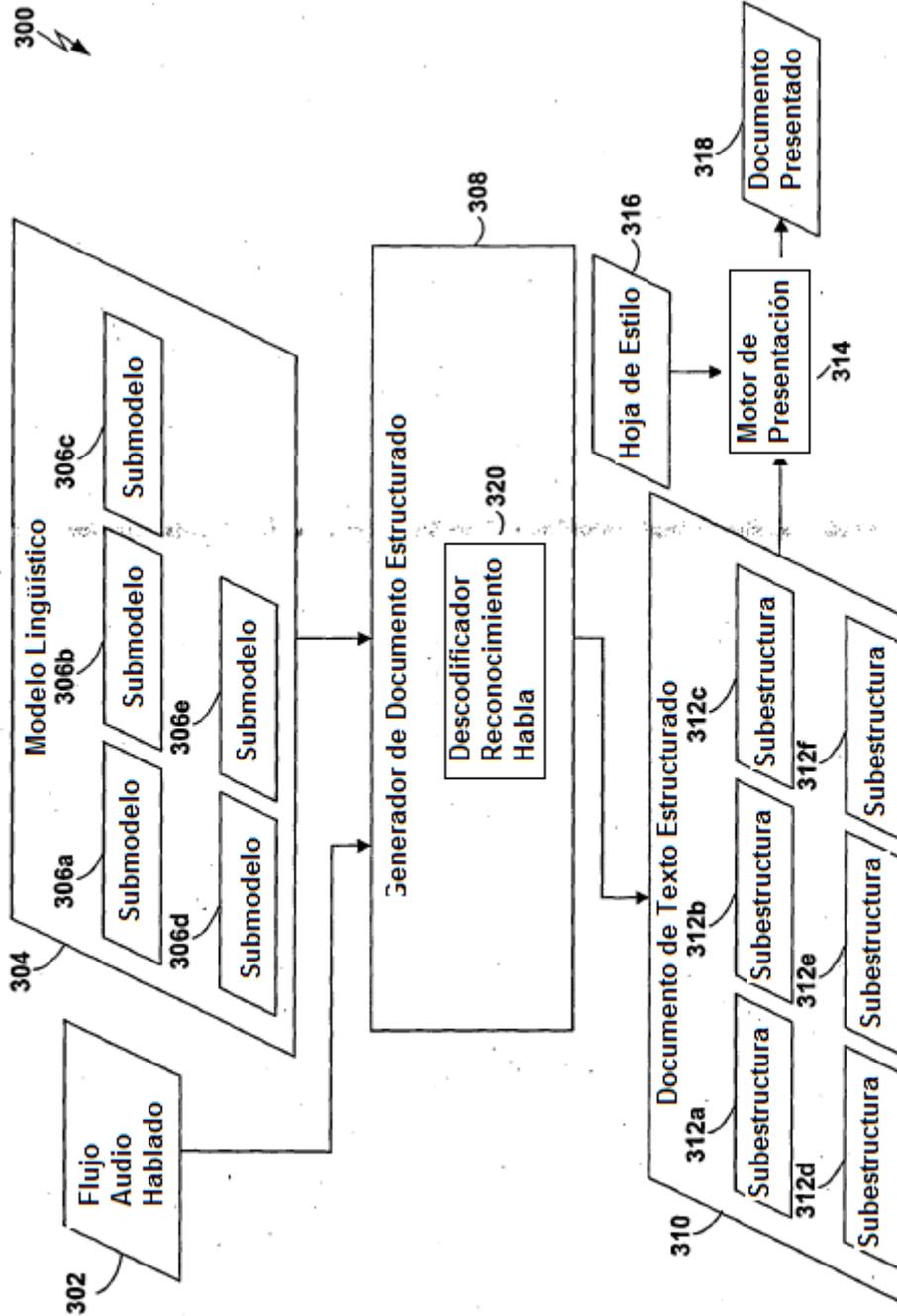
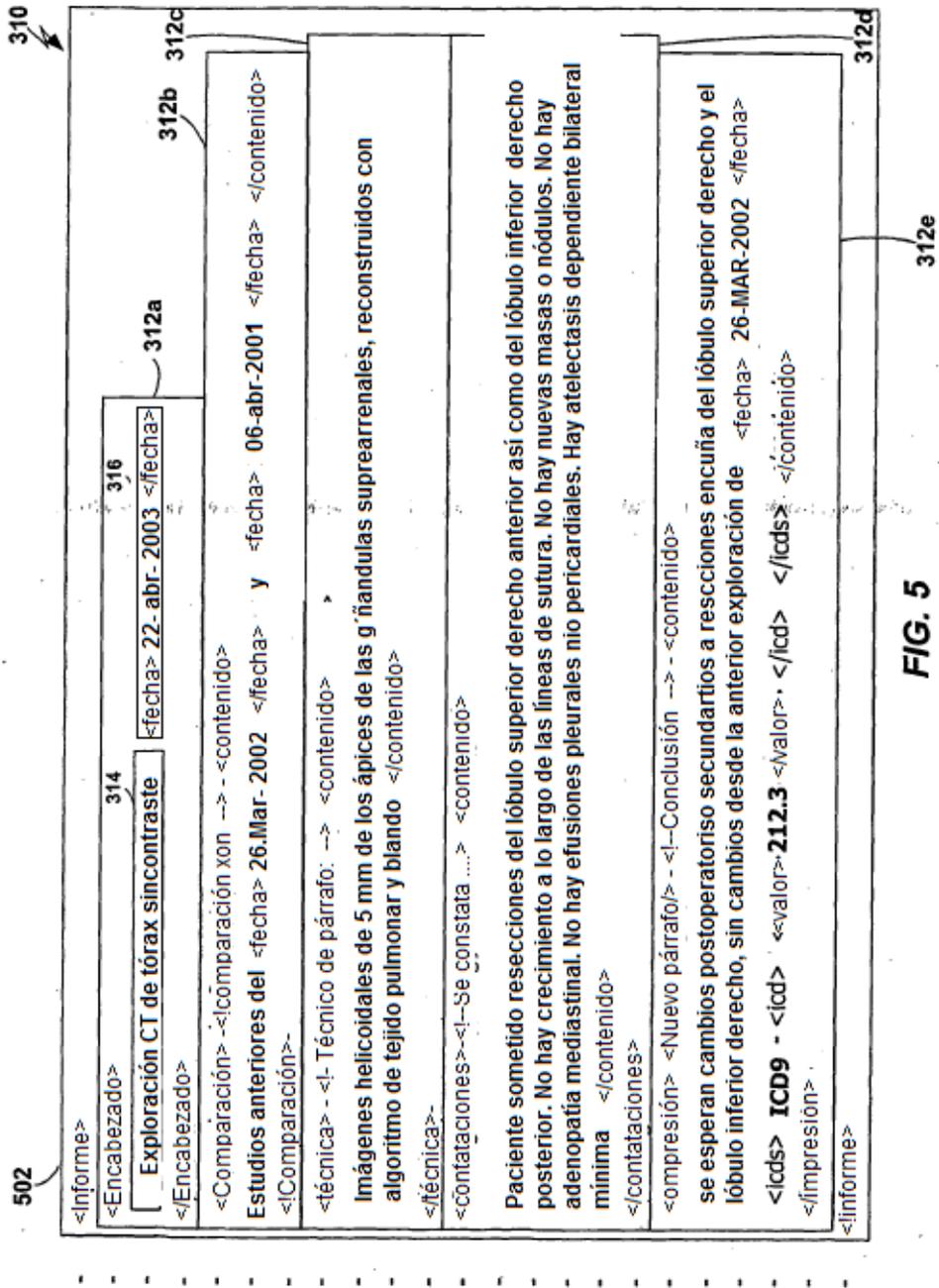


FIG. 3

302 

Exploración CT de tórax sin contraste veintidos de abril de dos mil tres comparaciones con estudios anteriores del veintiseis de marzo de dos mil dos y elseis de abril de dos mil uno, párrafo técnico %dos puntos% imágenes helicoidales de cin mm de los ápices a través de las glándulas suprarrenales %com% reconstruidos con el algoritmo de tejido pulmonar y blando %punto- paciente sometido a rsecciones en cuña del lóbulo superior derecho anterior asi como dellóbulo inferior derecho posterior %punto% no hay crecimiento a lo largo de las líneas de sutura %punto% no hay nuevas masas o nódulos %punto% no hay adenopatía mediastinal %punto% no hay efusiones pleurales ni pericárdicas %punto% hay atelectasis dependiente bilateral mínima %punto% %nuevo párrafo% conclusión se esperan cambios postoperatorios secundarios a las resecciones en cuña del lóbulo superior derecho y el lóbulo inferior derecho %coma% sin cambios desde la anterior exploración CT del veintiseis de marzo de dos mil dos % punto% ICD9 doscientos dos coma tres

Fig.4



<p>Exploración CT del Torax</p>	<p>602a</p>	<p>22-APR-2003</p>
<p>Comparación: Estudios anteriores del 26-MAR-2002 y 06-ABR-2001</p>	<p>602b</p>	
<p>Técnica: Imágenes helicoidales de 5 mm de los ápices a través de las glándulas suprarrenales, reconstruidos con algoritmo de tejido pulmonar y blando</p>	<p>602c</p>	
<p>Constataciones: Paciente sometido a resecciones en cuña del lóbulo superior derecho anterior así como del lóbulo inferior derecho posterior. No hay crecimiento a lo largo de las líneas de sutura. no hay nuevas masas o nódulos. No hay adenopatía mediastinal. No hay efusiones pleurales ni pericárdicas. Hay atelectasis dependiente bilateral mínima</p>	<p>602d</p>	
<p>Impresión: Se esperan cambios postoperatorios secundarios a las resecciones en cuña del lóbulo superior derecho y el lóbulo inferior derecho, sin cambios desde la anterior exploración CT del 26-MAR-2002.</p>	<p>602e</p>	<p>ICD9 212.3</p>

FIG. 6

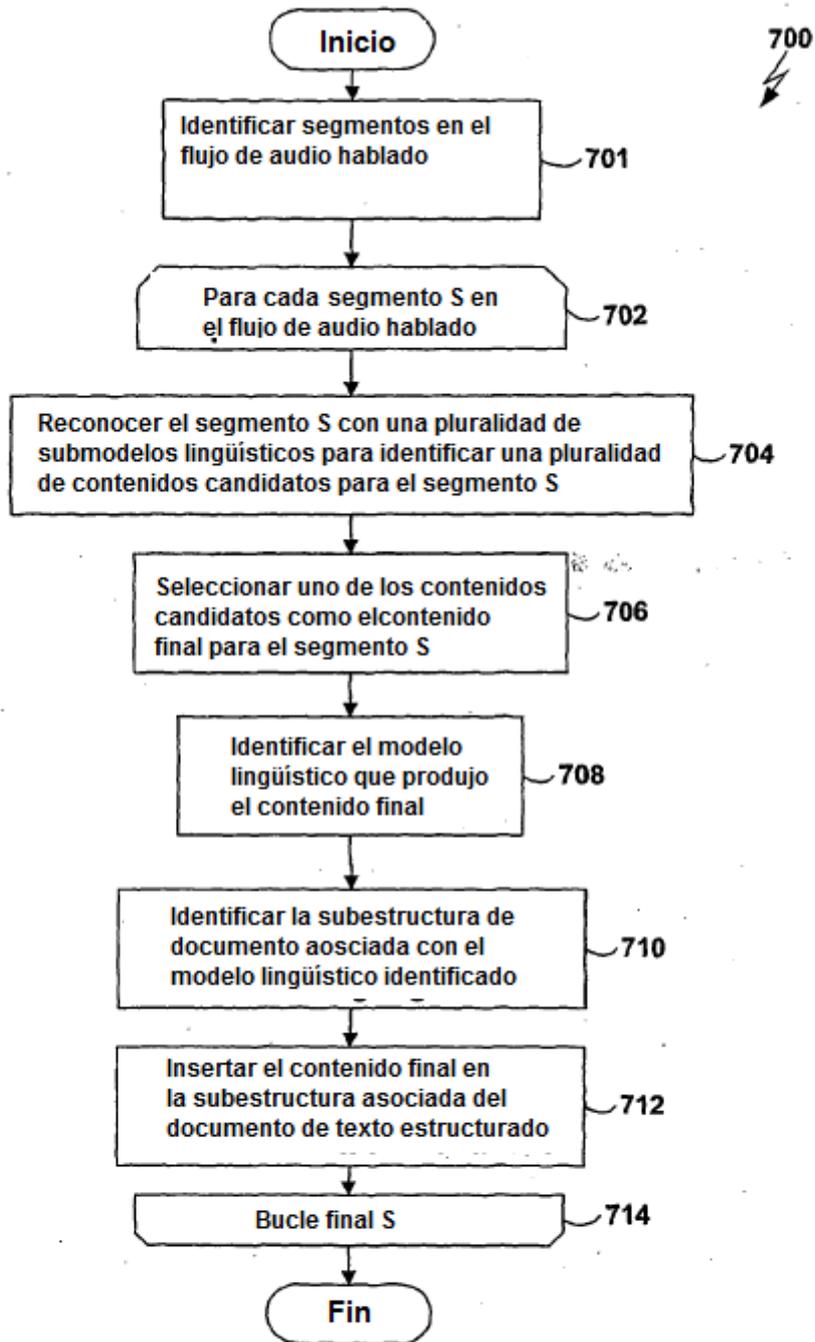


FIG. 7

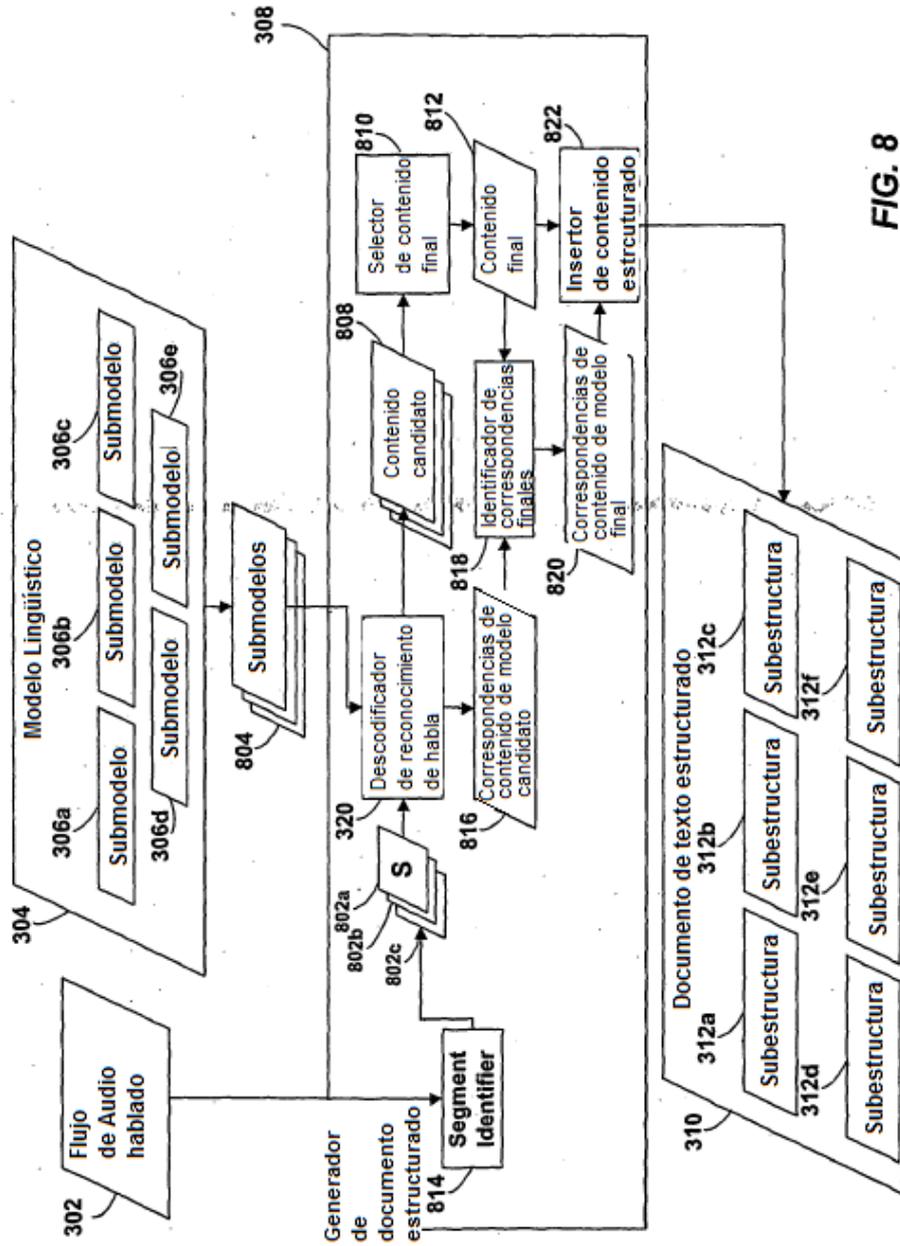


FIG. 8

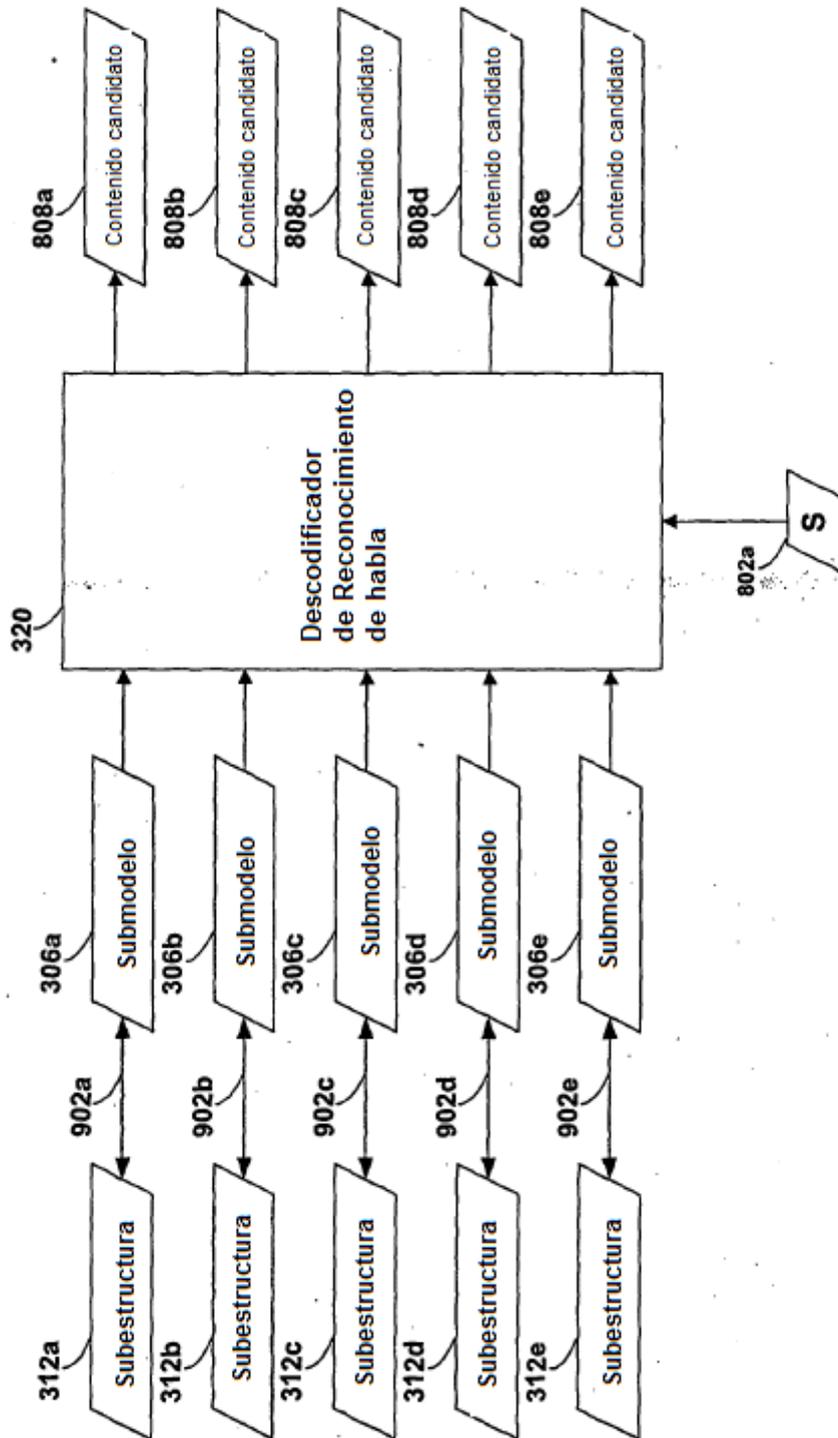


FIG. 9

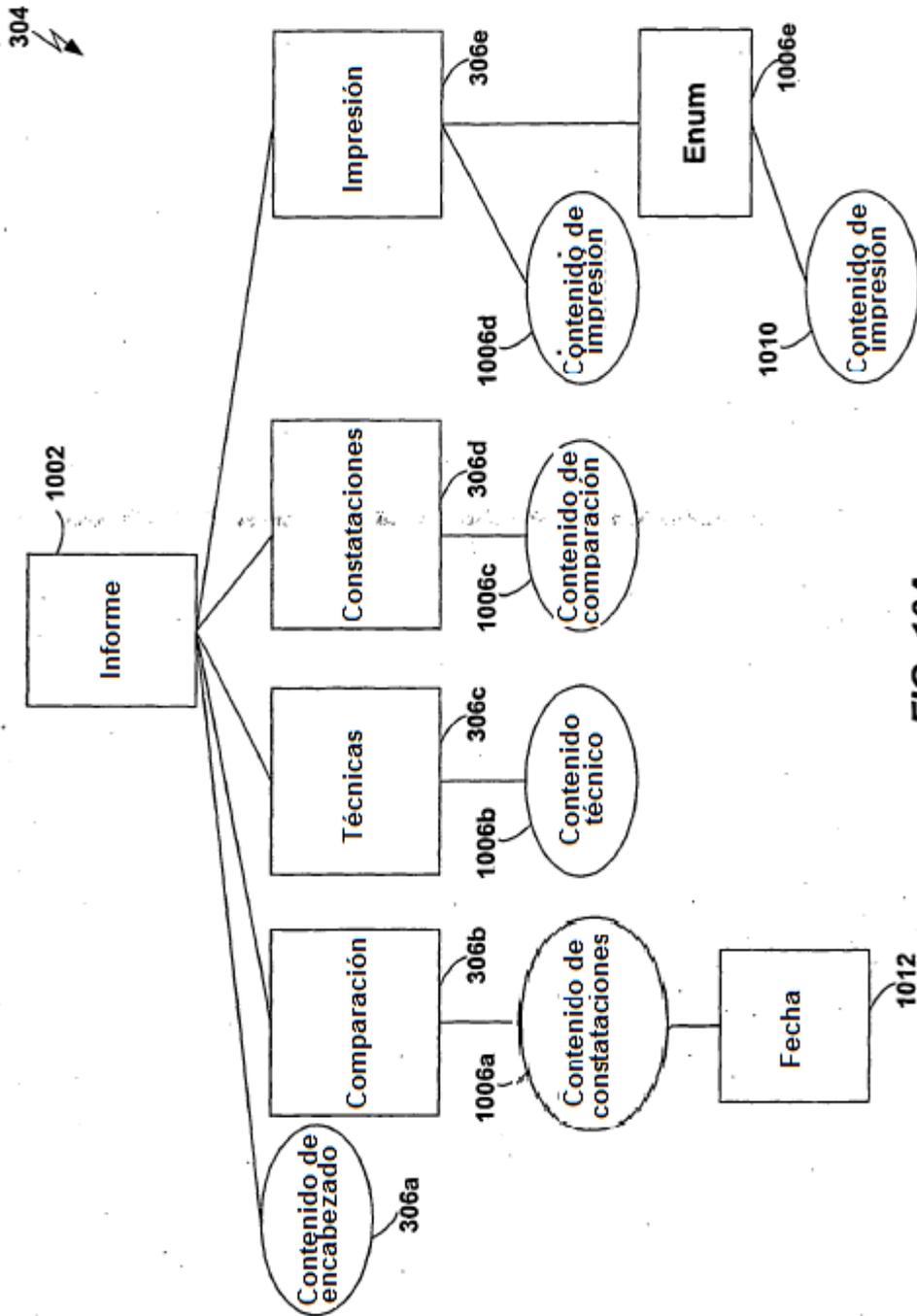


FIG. 10A

304

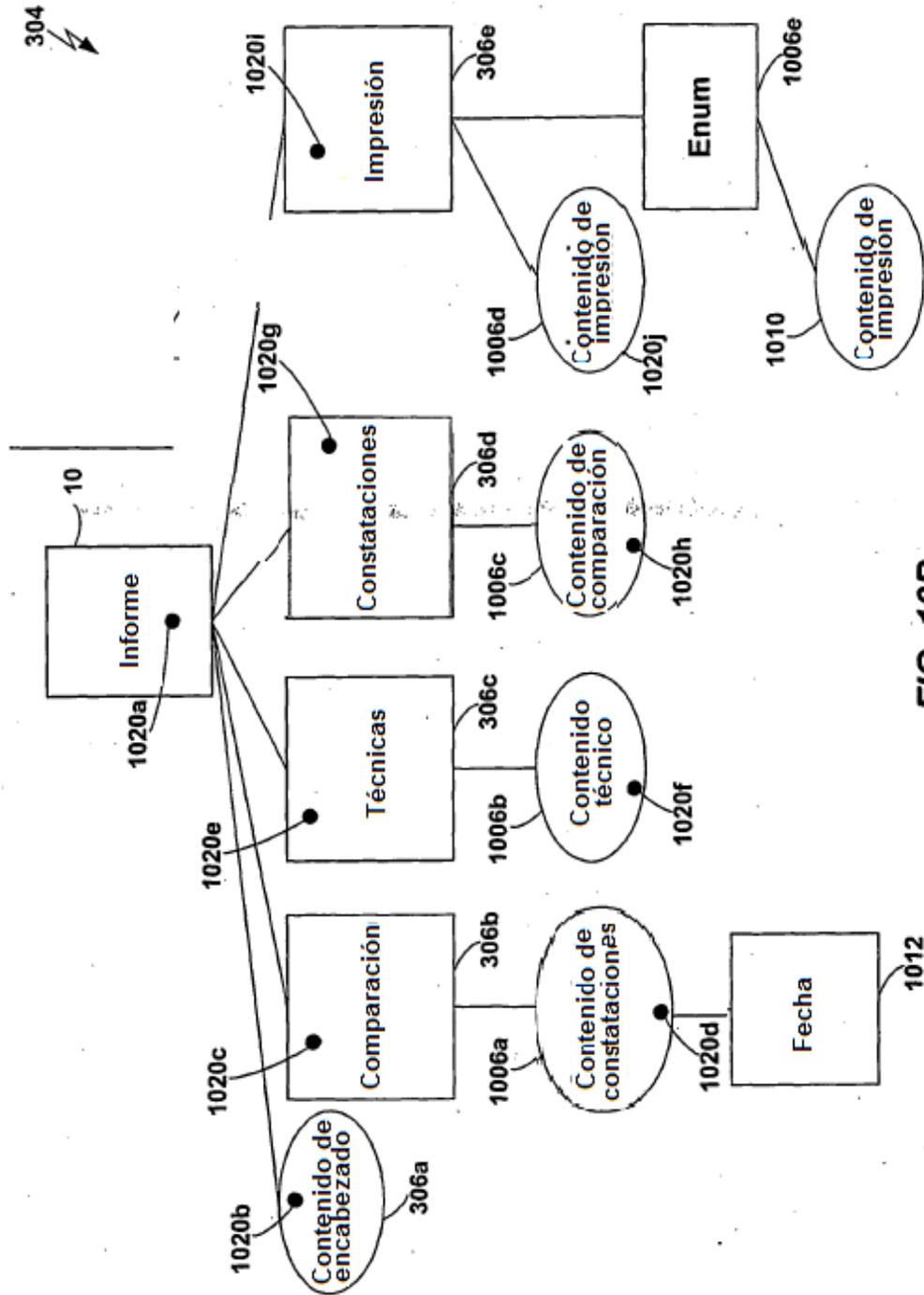


FIG. 10B

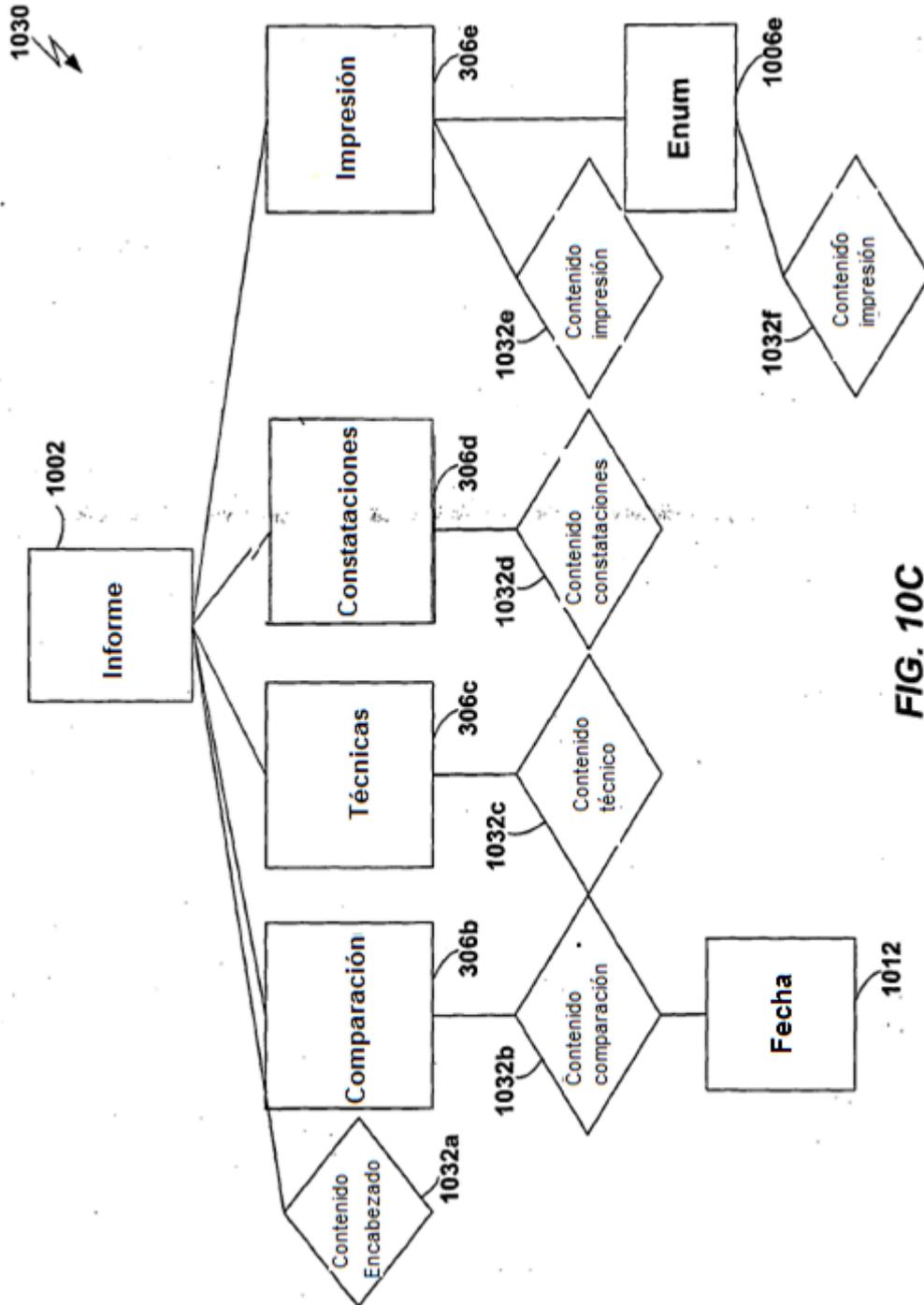


FIG. 10C

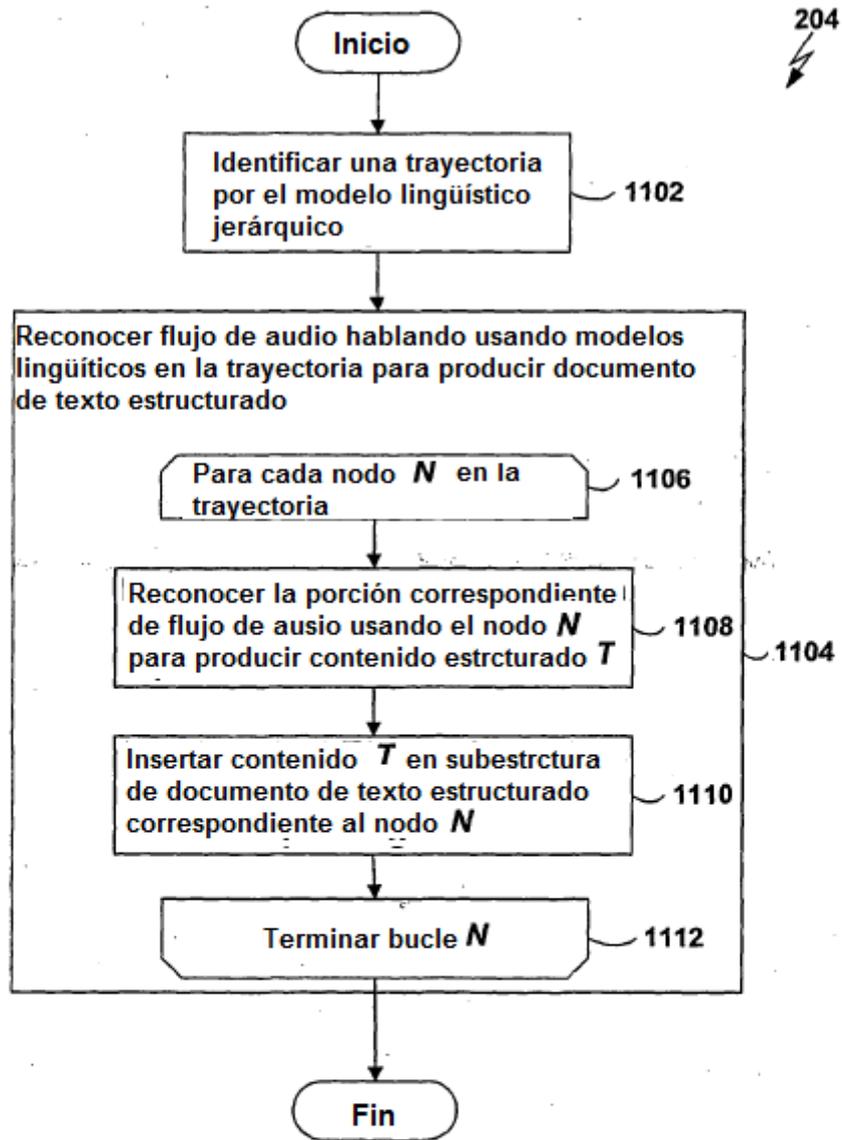


FIG. 11A

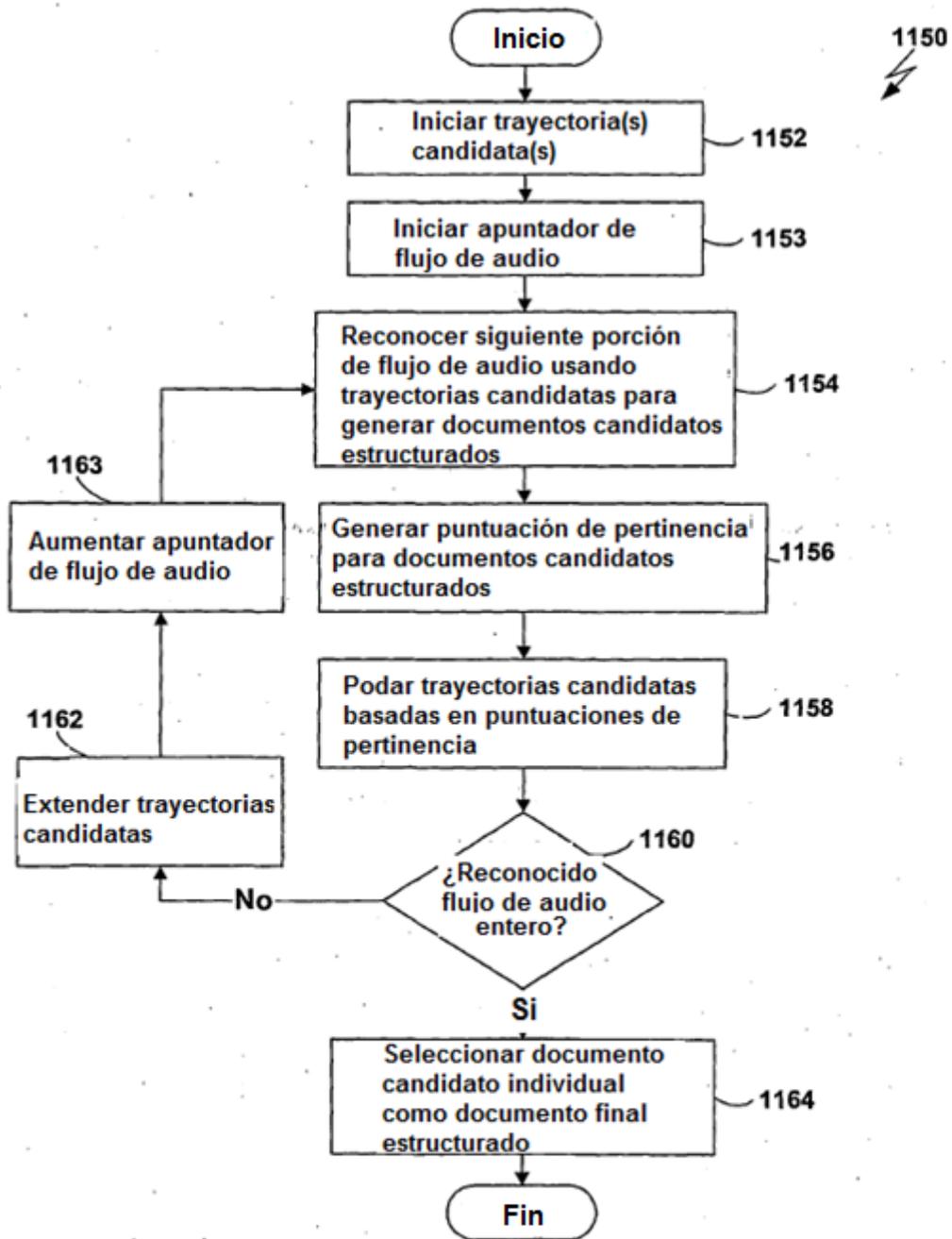


FIG. 11B

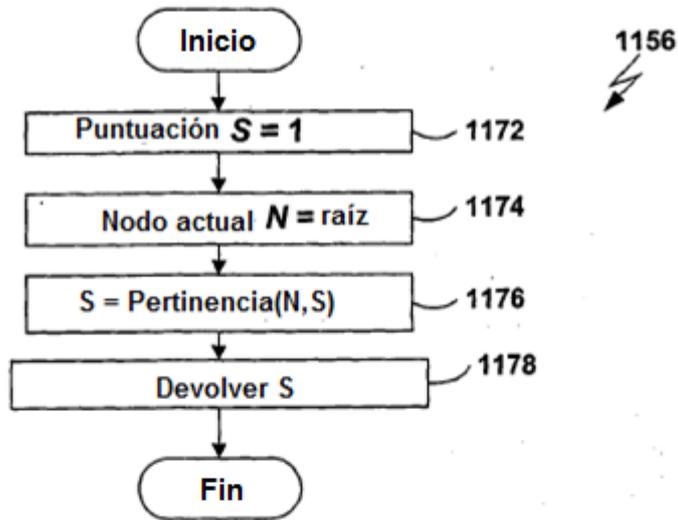


FIG. 11C

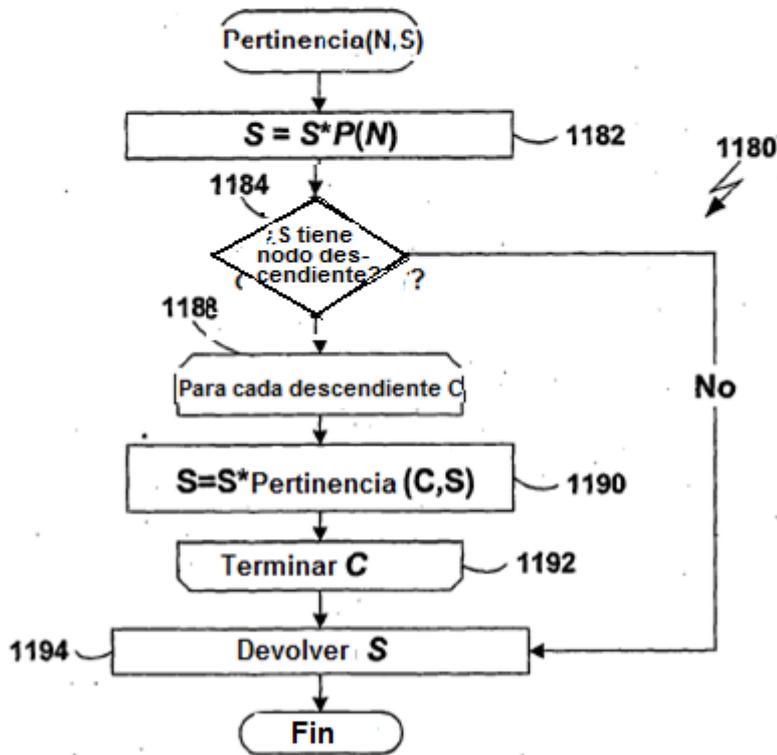


FIG. 11D

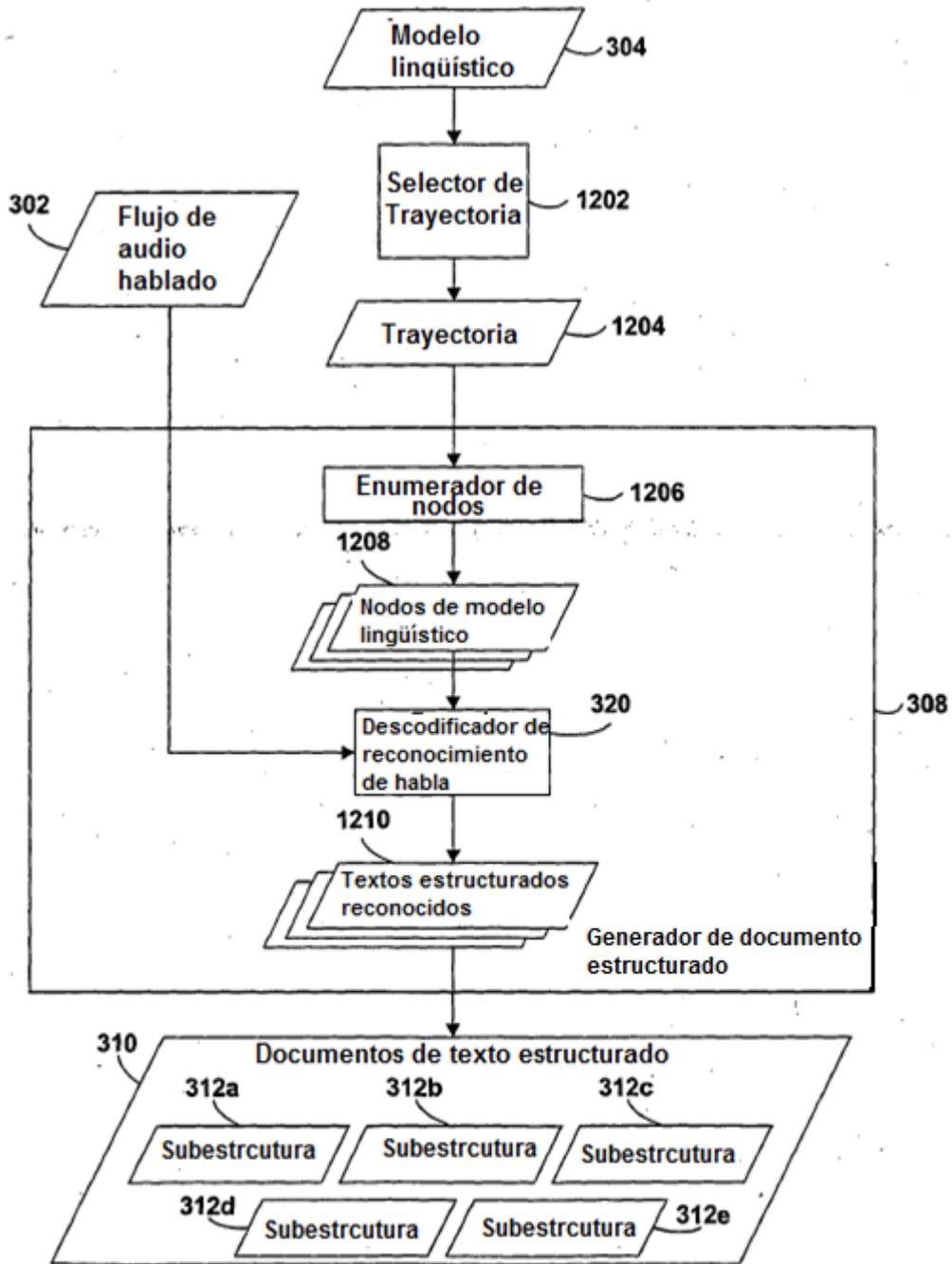


FIG. 12A

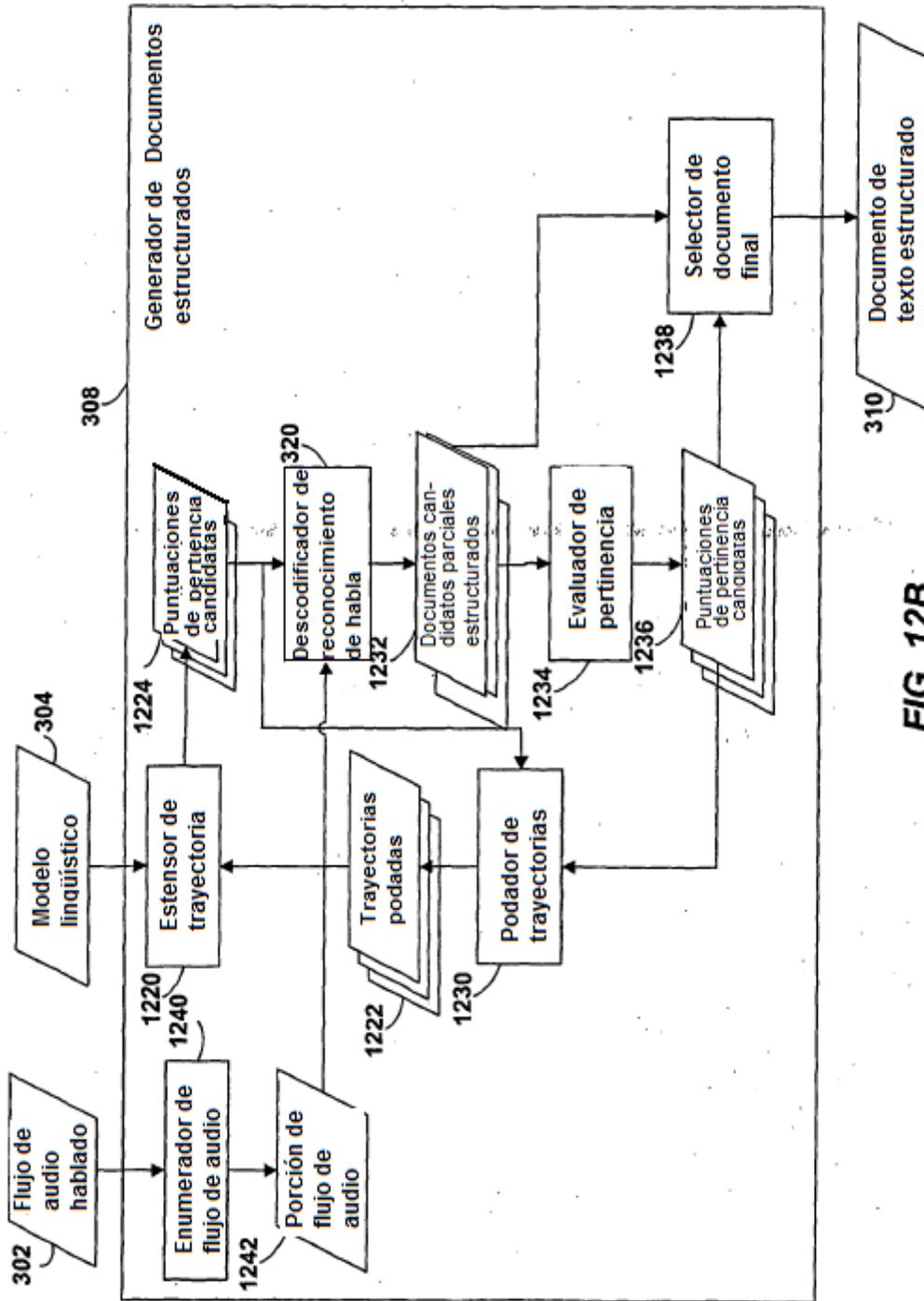


FIG. 12B

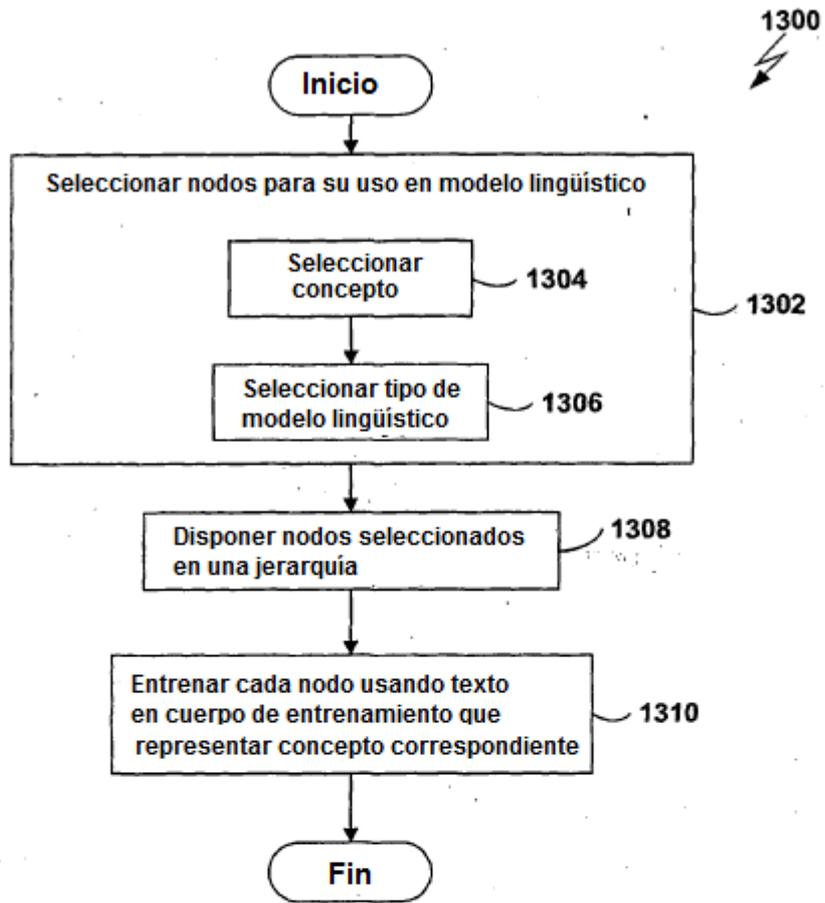


FIG. 13

1400
⚡

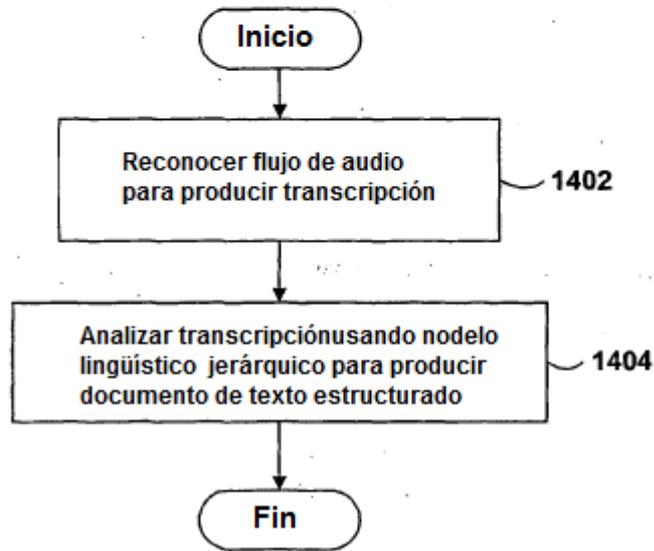


FIG. 14

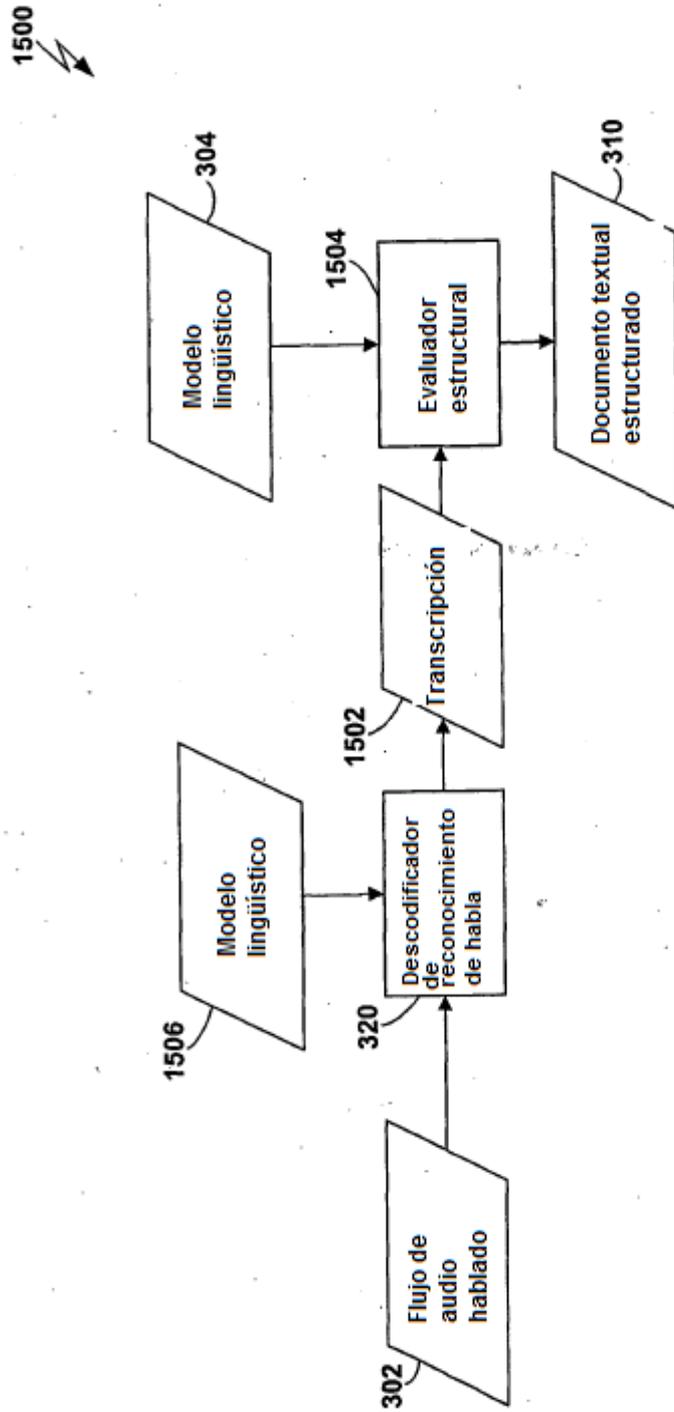


FIG. 15