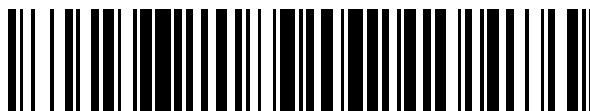


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 894 123**

51 Int. Cl.:

G10L 13/00 (2006.01)
G10L 13/10 (2013.01)
G10L 13/08 (2013.01)
H04N 21/2187 (2011.01)
H04N 21/2368 (2011.01)
H04N 21/488 (2011.01)
H04N 21/81 (2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **23.05.2018 PCT/JP2018/019828**
- 87 Fecha y número de publicación internacional: **29.11.2018 WO18216729**
- 96 Fecha de presentación y número de la solicitud europea: **23.05.2018 E 18805990 (1)**
- 97 Fecha y número de publicación de la concesión europea: **08.09.2021 EP 3633671**

54 Título: **Dispositivo de generación de orientación de audio, procedimiento de generación de orientación de audio y sistema de difusión**

30 Prioridad:

24.05.2017 JP 2017102847

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
11.02.2022

73 Titular/es:

NIPPON HOSO KYOKAI (50.0%)
2-1, Jinnan 2-chome Shibuya-ku
Tokyo 150-8001, JP y
NHK ENGINEERING SYSTEM, INC. (50.0%)

72 Inventor/es:

KUMANO TADASHI;
YAMADA ICHIRO;
IMAI ATSUSHI;
SUMIYOSHI HIDEKI;
YAMANOUCHI YUKO;
SHIMIZU TOSHIHIRO;
SEIYAMA NOBUMASA;
SATO SHOEI;
SAITO REIKO;
MIYAZAKI TARO;
KURIHARA KIYOSHI;
ICHIKI MANON y
TAKAGI TOHRU

74 Agente/Representante:

PONTI & PARTNERS, S.L.P.

ES 2 894 123 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

Dispositivo de generación de orientación de audio, procedimiento de generación de orientación de audio y sistema de difusión

5

CAMPO TÉCNICO

[0001] La presente invención se refiere a un dispositivo de generación de orientación de audio, un procedimiento de generación de orientación de audio y un sistema de difusión.

10

[0002] La prioridad se reivindica en la solicitud de patente japonesa n.º 2017-102847, depositada el 24 de mayo de 2017. ANTECEDENTES DE LA TÉCNICA

[0003] Al proporcionar un programa deportivo que incluye vídeo y audio de una competición deportiva en la difusión de televisión u otras formas de transmisión, se practica ampliamente la superposición de audio en vivo que describe la competición. Esto contribuye a que los programas deportivos sean más fáciles de entender y disfrutar.

[0004] En el arte relacionado, con el fin de producir audio en vivo, los locutores que están familiarizados con la competición siempre se han colocado en lugares de juego o similares para proporcionar comentarios en vivo durante la competición. Dado que este comentario en vivo es costoso, puede haber sido abandonado en muchos programas deportivos, o puede ser proporcionado en una forma que no superponga audio en vivo. Esto es especialmente cierto en eventos deportivos a gran escala donde muchas competiciones se celebran en paralelo.

[0005] Por lo tanto, se ha propuesto un procedimiento para generar automáticamente un texto que explique la situación de la competición a partir de los datos de competición de la competición deportiva. Por ejemplo, NPL 1 describe la generación automática de texto de ruptura de entrada en función de los resultados del bateador en cada entrada de béisbol, la información que se muestra en un marcador y similares. PL 1 describe un procedimiento de espectador que monitorea un estado de un evento en línea, actualizando un modelo de espectador, de modo que se puedan generar flujos de datos de espectador y proporcionar a los espectadores. El procedimiento de espectador puede generar comentarios virtuales para una acción ocurrida en el evento para generar los flujos de datos del espectador.

Lista de referencias

35 Bibliografía no relacionada con patentes

[0006] [NPL 1] Murakami Soichiro, Kanno Tienhei, Takamura Daiya, Okumura Manabu, «Automatic Generation of Inning Bulletin from Batter Performance», Language Processing Society 22nd Annual Conference, Proceedings, p. 338-341, 29 de febrero de 2016.

40

Bibliografía de patente

[0007] [PL 11] Publicación de solicitud de patente de los Estados Unidos, n.º: US 2006/058103 A1_

[0008] Por otro lado, incluso en un programa de difusión de televisión en el que se superpone el comentario en vivo hablado por un locutor o comentarista, cuando un oyente escucha el audio de la competición deportiva, a veces es imposible comprender suficientemente la situación de una competición solo por el comentario en vivo hablado por el locutor o comentarista. Además, en la difusión de subtítulos en vivo, hay un retraso de aproximadamente 10 segundos después de que se emite un comentario en vivo hasta que se presenta un subtítulo que indica el contenido del habla. Esto también puede ser causa de una comprensión insuficiente de la última situación de la competición.

[0009] Los organismos de difusión pueden emitir comentarios de manera que las personas con discapacidad visual puedan disfrutar más de los programas de difusión de televisión. La difusión de comentarios es un servicio de difusión que proporciona los contenidos que se muestran en la pantalla en la voz del presentador, como un locutor. En el arte relacionado, se necesita mucha gente para producir contenido para la difusión de comentarios. Por lo tanto, la proporción de programas de difusión en los que se difunden comentarios es limitada entre los programas de difusión proporcionados por las emisoras. Además, se han producido subtítulos mediante procesamiento de reconocimiento de voz u operación manual basada en voces habladas por locutores en vivo y similares. Para dicho contenido del habla, se ha considerado que proporciona información complementaria similar a la del comentario difundido por voz, utilizando la técnica descrita en la NPL 1 descrita anteriormente.

60

DESCRIPCIÓN DE LA INVENCION

Problema técnico

[0010] Sin embargo, el procedimiento descrito en NPL 1 se limita al béisbol, y es un procedimiento para analizar juntos los datos para cada entrada y generar texto explicativo que resume la situación que ocurre en cada entrada. Es decir, el procedimiento descrito en la NPL 1 no tiene en cuenta la generación de texto explicativo junto con un vídeo de difusión en vivo como una difusión deportiva.

5

[0011] La presente invención se ha realizado en vista de los puntos anteriores, y un objetivo de la misma es proporcionar un dispositivo de generación de orientación de audio, un dispositivo de generación de orientación de audio, un procedimiento de generación de orientación de audio y un sistema de difusión, capaz de transmitir con precisión la situación de una competición junto con el vídeo.

10

Solución del problema

[0012] La presente invención se ha realizado para resolver los problemas descritos anteriormente, y un aspecto de la presente invención es un dispositivo de generación de orientación de audio que comprende una unidad de gestión de mensajes que recibe y acumula un mensaje, en el que el mensaje se distribuye para cada actualización, el mensaje son datos que representan una situación más reciente de una competición; una unidad de generación de explicación que genera un texto explicativo para transmitir información no transmitida detectada a partir del mensaje, en función de la información transmitida; y una unidad de síntesis de habla que produce un habla convertida a partir del texto explicativo, en la que la unidad de generación de explicación almacena la información no transmitida para el texto explicativo, mientras que la información transmitida después de generar el texto explicativo, permanece en espera hasta que se completa la salida del habla e inicia un procedimiento para generar un nuevo texto explicativo en función de información no transmitida actualizada.

15

20

[0013] Además, el dispositivo de generación de orientación de audio se caracteriza porque la unidad de gestión de mensajes detecta una porción modificada del mensaje debido a la actualización del mensaje, genera nueva información según un tipo de la parte variada y acumula adicionalmente la nueva información generada.

25

[0014] Además, en el dispositivo de generación de orientación de audio, la unidad de gestión de mensajes puede seleccionar una plantilla de generación de información correspondiente al tipo de la porción modificada a partir de plantillas de generación de información prealmacenadas, cada una de las plantillas de generación de información prealmacenadas indica una estructura de la nueva información para cada tipo de la porción modificada, y generar la nueva información mediante la integración de elementos de la parte variada en la plantilla de generación de información seleccionada.

30

[0015] Además, en el dispositivo de generación de orientación de audio, la unidad de generación de explicación puede considerar que los tipos de información predeterminados, entre las piezas de información retenidas como la información transmitida, no se han transmitido, en un caso en el que ha transcurrido un tiempo predeterminado desde un tiempo en que se completa la transmisión.

35

[0016] Además, en el dispositivo de generación de orientación de audio, la unidad de generación de explicación puede seleccionar una plantilla explicativa correspondiente al tipo de la información no transmitida a partir de plantillas explicativas prealmacenadas, cada una de las plantillas explicativas prealmacenadas indica un patrón de oración del texto explicativo para cada tipo de información no transmitida, y generar el texto explicativo mediante la integración de elementos de la información no transmitida en la plantilla explicativa seleccionada.

40

45

[0017] En el dispositivo de generación de orientación de audio, la unidad de generación de explicación puede utilizar una plantilla de texto en la que una porción acentuada está predeterminada para los datos de la competición que indican una situación de la competición indicada por el mensaje para generar el texto explicativo que indica la porción acentuada de los datos de la competición, la unidad de síntesis de habla puede comprender una unidad de generación de rasgo de lenguaje de fonema que analiza el texto explicativo para generar un rasgo de lenguaje para cada fonema e información de entonación acentuada para el fonema; una unidad de cálculo de duración del tiempo que genera una duración del tiempo del fonema a partir del rasgo de lenguaje y la información de entonación acentuada generada por la unidad de generación de rasgo de lenguaje de fonema, utilizando un modelo de duración del tiempo; una unidad de generación de rasgo de lenguaje de trama que genera un rasgo de lenguaje para cada trama del fonema a partir de la duración del tiempo del fonema generado por la unidad de cálculo de duración del tiempo y el rasgo de lenguaje del fonema, y asocia el rasgo de lenguaje generado con la información de entonación acentuada del fonema correspondiente a la trama; una unidad de cálculo de rasgo acústico que genera un rasgo acústico para cada trama, a partir del rasgo de lenguaje de la trama generada por la unidad de generación de rasgo de lenguaje de trama y la información de entonación acentuada correspondiente a la trama, utilizando un modelo de rasgo acústico; y una unidad de síntesis que sintetiza el habla usando el rasgo acústico para cada trama generada por la unidad de cálculo de rasgo acústico, el modelo de duración del tiempo puede aprenderse preliminarmente, usando el rasgo de lenguaje para cada fonema, e información de entonación acentuada que indica la presencia o ausencia de entonación acentuada para el fonema como una entrada, y la duración del tiempo del fonema como una salida, y el modelo de rasgo acústico puede aprenderse preliminarmente, usando el rasgo de lenguaje para cada trama del fonema e información de entonación acentuada que indica la presencia o ausencia de entonación acentuada de la trama como una entrada, y la duración

50

55

60

65

del tiempo del fonema como una salida, y un rasgo acústico de la trama como una salida.

[0018] En el dispositivo de generación de orientación de audio, el modelo de duración del tiempo y el modelo de rasgo acústico se pueden aprender adicionalmente usando información de terminación de oración que indica si el fonema es un fonema de una palabra de terminación de oración que tiene una parte del habla como un sustantivo, la unidad de generación de rasgo de lenguaje de fonema puede generar la información de terminación de oración para cada fonema, en función de la parte del habla de la palabra de terminación de oración del texto, la unidad de cálculo de duración del tiempo puede generar una duración del tiempo del fonema, mediante el uso del modelo de duración del tiempo, a partir del rasgo de lenguaje del fonema, la información de entonación acentuada y la información de terminación de oración, la unidad de generación de rasgo de lenguaje de trama puede asociar el rasgo de lenguaje para cada trama con la información de terminación de oración, y la unidad de cálculo de rasgo acústico puede generar un rasgo acústico de cada trama, mediante el uso del modelo de rasgo acústico, a partir del rasgo de lenguaje de la trama, y la entonación acentuada y la información de terminación de oración correspondiente a la trama.

[0019] Otro aspecto de la presente invención se refiere a un sistema de difusión que incluye el dispositivo de generación de orientación de audio según la reivindicación 1; una primera unidad de búfer que almacena datos de audio que indican el habla convertida por la unidad de síntesis de habla e información de tiempo que indica el tiempo de una situación de la competición utilizada para generar el texto explicativo; una segunda unidad de búfer que almacena datos de vídeo que indican la situación de la competición en asociación con información de tiempo que indica la hora de la situación; y una unidad de multiplexación que extrae datos de audio en un momento designado de la primera unidad de búfer, el tiempo designado es una primera duración predeterminada antes de un tiempo actual, extrae datos de vídeo en el momento designado de la segunda unidad de búfer y multiplexa los datos de audio y los datos de vídeo.

[0020] Aún otro aspecto de la presente invención puede ser el sistema de difusión descrito anteriormente, en el que la unidad de gestión de mensajes almacena los datos de competición que incluyen el mensaje en asociación con información de tiempo que indica la hora de la situación de la competición indicada por el mensaje, la unidad de generación de explicación se refiere a los datos de competición en un momento como información no transmitida de la unidad de gestión de mensajes, siendo el tiempo una segunda duración predeterminada anterior al tiempo actual, y la segunda duración predeterminada es más corta que la duración predeterminada.

[0021] Además, otro aspecto de la presente invención se refiere a un sistema de difusión que incluye el dispositivo de generación de orientación de audio según la reivindicación 1; y un dispositivo transmisor que transmite el texto explicativo y el habla.

[0022] Otro aspecto adicional de la presente invención es un procedimiento de generación de orientación de audio según la reivindicación 10.

Efectos ventajosos de la invención

[0023] Según la presente invención, es posible proporcionar audio que transmita con precisión la situación de una competición junto con el vídeo.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0024]

La Fig. 1 es un diagrama de bloques que ilustra un ejemplo de configuración de un sistema de difusión según una primera realización.

La Fig. 2 es un diagrama de bloques que ilustra un ejemplo de configuración de un dispositivo de generación de orientación de audio según la primera realización.

La Fig. 3 es un diagrama que muestra un ejemplo de un mensaje según la primera realización.

La Fig. 4 es un diagrama que muestra otro ejemplo del mensaje según la primera realización.

La Fig. 5 es un diagrama de flujo que ilustra un primer ejemplo de procesamiento de gestión de mensajes según la primera realización.

La Fig. 6 es un diagrama de flujo que ilustra un primer ejemplo de procesamiento de gestión de mensajes según una segunda realización.

La Fig. 7 es un diagrama de flujo que ilustra un primer ejemplo de procesamiento de gestión de mensajes según una tercera realización.

La Fig. 8 es un diagrama explicativo que muestra un ejemplo de procesamiento de generación de información de eventos según la primera realización.

La Fig. 9 es un diagrama de flujo que ilustra un ejemplo de procesamiento de generación de explicación según la primera realización.

La Fig. 10 es un diagrama que muestra un ejemplo de una plantilla explicativa según la primera realización.

La Fig. 11 es un diagrama que muestra un ejemplo de una regla de selección de plantilla explicativa según la

primera realización.

La Fig. 12 es un diagrama que muestra un ejemplo de un texto explicativo según la primera realización.

La Fig. 13 es un diagrama de bloques que ilustra un ejemplo de configuración de un dispositivo de generación de orientación de audio según una segunda realización.

5 La Fig. 14 es un diagrama de bloques que ilustra un ejemplo de configuración de una unidad de preaprendizaje según la segunda realización.

La Fig. 15 es un diagrama de bloques que ilustra un ejemplo de configuración de una unidad de síntesis de habla según la segunda realización.

10 La Fig. 16 es un diagrama que muestra un ejemplo de un modelo de duración del tiempo según la segunda realización.

La Fig. 17 es un diagrama que muestra un ejemplo de un modelo de rasgo acústico según la segunda realización.

La Fig. 18 es un diagrama que muestra un ejemplo de datos de competición según la segunda realización.

La Fig. 19 es un diagrama que muestra otro ejemplo de los datos de competición según la segunda realización.

La Fig. 20 es un diagrama que muestra un ejemplo de una plantilla explicativa según la segunda realización.

15 La Fig. 21 es un diagrama que muestra otro ejemplo de la plantilla explicativa según la segunda realización.

La Fig. 22 es un diagrama que muestra aún otro ejemplo de la plantilla explicativa según la segunda realización.

La Fig. 23 es un diagrama que muestra un ejemplo de un texto generado según la segunda realización.

La Fig. 24 es un diagrama que muestra otro ejemplo de un texto generado según la segunda realización.

20 La Fig. 25 es un diagrama que muestra un ejemplo de entonación del habla sintetizada generada según la segunda realización.

La Fig. 26 es un diagrama que muestra otro ejemplo de entonación del habla sintetizada generada según la segunda realización.

La Fig. 27 es un diagrama de flujo que ilustra un ejemplo de preaprendizaje según la segunda realización.

25 La Fig. 28 es un diagrama de flujo que ilustra un ejemplo de procesamiento de generación de descripción de audio según la segunda realización.

La Fig. 29 es un diagrama de bloques que ilustra un ejemplo de configuración de un dispositivo de generación de orientación de audio según una tercera realización.

La Fig. 30 es un diagrama de flujo que ilustra un ejemplo de procesamiento de distribución de descripción de audio según la tercera realización.

30 MEJOR FORMA DE LLEVAR A CABO LA PRESENTE INVENCION <Primera realización

[0025] En lo sucesivo, se describirá en detalle la primera realización ejemplar de la presente invención con referencia a los dibujos adjuntos.

35 [0026] La Fig. 1 es un diagrama de bloques que ilustra un ejemplo de configuración de un sistema de difusión 1 según la primera realización. El sistema de difusión 1 incluye un dispositivo de generación de orientación de audio 10, una unidad de procesamiento de vídeo 15, un dispositivo transmisor 16 y un dispositivo receptor 20. El dispositivo de generación de orientación de audio 10, la unidad de procesamiento de vídeo 15 y el dispositivo transmisor 16 son un grupo de dispositivos del lado de transmisión que pertenece principalmente a una emisora. El dispositivo receptor 20 es un dispositivo del lado de recepción que pertenece principalmente al espectador. En la siguiente descripción, se toma como ejemplo un caso en el que el sistema de difusión 1 se aplica a la difusión de televisión.

45 [0027] El dispositivo de generación de orientación de audio 10 obtiene la última versión de un mensaje, que son datos que representan la información más reciente sobre la competición, cada vez que se actualizan los datos. El dispositivo de generación de orientación de audio 10 recibe, como un mensaje, por ejemplo, datos de competición distribuidos en tiempo real desde un centro de datos a través de una red de comunicación. El centro de datos puede ser un centro de datos oficial para competiciones como Olympic Broadcasting Services (OBS). El dispositivo de generación de orientación de audio 10 puede recibir datos de competición directa o indirectamente de un dispositivo de edición de una emisora o un creador de contenido. Los datos de la competición son datos en un formato predeterminado (por ejemplo, Olympic Data Feed (ODF)) que indican la situación de la competición deportiva representada en el vídeo proporcionado por difusión o similar. Más adelante se describirá un ejemplo de datos sobre la competición.

55 [0028] El dispositivo de generación de orientación de audio 10 acumula el último mensaje de recepción adquirido. El dispositivo de generación de orientación de audio 10 detecta información no transmitida del mensaje de recepción en función de la información transmitida. El dispositivo de generación de orientación de audio 10 genera un texto explicativo para transmitir la información no transmitida detectada y emite los datos de audio convertidos del texto explicativo generado al dispositivo transmisor 16. El dispositivo de generación de orientación de audio 10 acumula información no transmitida utilizada para generar el texto explicativo como información transmitida. El dispositivo de generación de orientación de audio 10 espera solo el tiempo necesario para reproducir los datos de audio, y a continuación repite una serie de procedimientos para detectar nueva información no transmitida y generar un texto explicativo de la información no transmitida detectada. El dispositivo de generación de orientación de audio 10 puede emitir el texto explicativo generado al dispositivo transmisor 16.

65

- [0029]** La unidad de procesamiento de vídeo 15 adquiere datos de vídeo y datos de audio. Los datos de vídeo y audio adquiridos son datos que representan el vídeo y el audio de una competición deportiva indicados por los datos de la competición. La unidad de procesamiento de vídeo 15 recibe, por ejemplo, datos de vídeo de un dispositivo de imagen (por ejemplo, una cámara de vídeo) instalado en un lugar de competición, y datos de audio de un dispositivo de recolección de sonido (por ejemplo, un micrófono) directa o indirectamente. La unidad de procesamiento de vídeo 15 puede editar los datos de vídeo y audio de la sección o parte emitida, a partir de los datos de vídeo y audio adquiridos, según el funcionamiento del operador. La unidad de procesamiento de vídeo 15 emite los datos de vídeo y audio adquiridos o editados al dispositivo transmisor 16.
- 10 **[0030]** En la siguiente descripción, los datos de audio proporcionados desde la unidad de procesamiento de vídeo 15 se denominan datos de audio originales, los datos de audio generados por el dispositivo de generación de orientación de audio 10 se denominan datos de descripción de audio y se distinguen ambos datos de audio. Los datos de audio originales pueden ser solo el sonido de fondo grabado en el lugar, o pueden ser datos que indican que un habla, como un comentario en vivo, está mezclada con el sonido de fondo.
- 15 **[0031]** El dispositivo transmisor 16 multiplexa la entrada de datos de descripción de audio desde el dispositivo de generación de orientación de audio 10, y los datos de vídeo y los datos de audio originales que se introducen desde la unidad de procesamiento de vídeo 15. El dispositivo transmisor 16 transmite los datos multiplexados obtenidos por multiplexación a la vía de transmisión de difusión. La vía de transmisión de difusión es, por ejemplo, una onda de 20 difusión en una banda de frecuencia predeterminada. Una parte de la vía de transmisión de difusión puede incluir una línea de comunicación.
- [0032]** Los datos de descripción de audio y los datos de audio originales se pueden multiplexar como subvoz y voz principal en la difusión multiplexada de audio, respectivamente. Cuando el texto explicativo se introduce desde el 25 dispositivo de generación de orientación de audio 10, el dispositivo transmisor 16 puede además multiplexar el texto explicativo de entrada como datos de subtítulos.
- [0033]** El dispositivo receptor 20 recibe datos multiplexados transmitidos a través de la vía de transmisión de difusión, y presenta el contenido del programa de difusión en función de los datos multiplexados recibidos. El 30 dispositivo receptor 20 separa los datos multiplexados recibidos en datos de vídeo, datos de audio originales y datos de descripción de audio. El dispositivo receptor 20 incluye una unidad de visualización (por ejemplo, una pantalla) y una unidad de reproducción (por ejemplo, un altavoz). La unidad de visualización muestra un vídeo en función de los datos de vídeo separados. La unidad de reproducción reproduce audio relacionado con uno o ambos de los datos de audio originales y los datos de descripción de audio. El dispositivo receptor 20 puede seleccionar uno del audio en 35 función de los datos de audio originales y el audio de orientación en función de los datos de descripción de audio, según una operación del usuario (por ejemplo, presionar el botón de voz principal o el botón de subvoz del controlador remoto). El dispositivo receptor 20 puede mezclar los datos de audio originales y el audio de orientación, y hacer que la unidad de reproducción reproduzca el habla mixta obtenida por la mezcla.
- 40 **[0034]** En un caso en el que los datos multiplexados incluyen datos de subtítulos, el dispositivo receptor 20 puede separar los datos de subtítulos de los datos multiplexados. El dispositivo receptor 20 puede superponer la cadena de caracteres indicada por los datos de subtítulos separados en el vídeo, en respuesta a una operación del usuario (por ejemplo, presionar un botón de subtítulos en el controlador remoto). El dispositivo receptor 20 muestra, en la unidad de visualización, un vídeo en el que se superpone la cadena de caracteres. El dispositivo receptor 20 es, 45 por ejemplo, un receptor de televisión.
- [0035]** El dispositivo de generación de orientación de audio 10 puede emitir el texto explicativo generado, a la unidad de procesamiento de vídeo 15, en lugar del dispositivo transmisor 16.
- 50 **[0036]** Cuando el texto explicativo se introduce desde el dispositivo de generación de orientación de audio 10, la unidad de procesamiento de vídeo 15 puede superponer la cadena de caracteres representada por el texto explicativo en los datos de vídeo como un subtítulo, y emitir los datos de vídeo en los que se superpone el subtítulo, al dispositivo transmisor 16.
- 55 (Dispositivo de generación de orientación de audio)
- [0037]** A continuación, se describirá un ejemplo de configuración del dispositivo de generación de orientación de audio 10 según la presente realización.
- 60 **[0038]** La Fig. 2 es un diagrama de bloques que ilustra un ejemplo de configuración del dispositivo de generación de orientación de audio 10 según la presente realización.
- [0039]** El dispositivo de generación de orientación de audio 10 incluye una unidad de gestión de mensajes 11, una unidad de generación de explicación 12 y una unidad de síntesis de habla 13.
- 65

[0040] La unidad de gestión de mensajes 11 contiene la última versión del mensaje de recepción adquirido. La unidad de gestión de mensajes 11 puede configurarse para contener la última versión para cada tipo, suponiendo que el mensaje de recepción adquirido se divide en una pluralidad de tipos diferentes. El tipo de mensaje es, por ejemplo, un mensaje de estado del juego que indica el estado de una competición (estado del juego), un mensaje de lista de eventos que enumera los eventos que han ocurrido, o similares. La unidad de gestión de mensajes 11 también mantiene el mensaje de recepción adquirido de la última versión junto con el mensaje antes de la última versión, y detecta una parte variada de mensaje comparando la versión anterior con un nuevo mensaje de recepción cuando se adquiere el nuevo mensaje de recepción. La unidad de gestión de mensajes 11 puede configurarse para generar nueva información según el tipo de variación detectado y a continuación agregar la información generada a un mensaje de un tipo de mensaje especial retenido por la unidad de gestión de mensajes 11. Como ejemplo, esta función se puede utilizar para generar información de eventos al detectar un evento a partir de una variación en el estado del juego en una competición para la que solo el estado del juego se distribuye como un mensaje, y para mantenerlo como una lista de eventos que incluye la información de eventos generada. El procesamiento realizado por la unidad de gestión de mensajes 11 se describirá más adelante.

[0041] Cuando la unidad de gestión de mensajes genera nueva información según el procedimiento anterior, la unidad de gestión de mensajes 11 puede utilizar una plantilla de generación de información almacenada previamente en la unidad de gestión de mensajes 11. La plantilla de generación de información son datos que indican una plantilla de una estructura de información para generar, como parte de un mensaje, nueva información que incluye un elemento de información, que tiene el elemento de información que es una parte predeterminada del mensaje como un espacio en blanco. La unidad de gestión de mensajes 11 determina si hay una plantilla de generación de información que coincide con una regla de selección predeterminada en función del tipo de parte variada del mensaje de recepción, de entre las plantillas de generación de información almacenadas, y selecciona la plantilla de generación de información correspondiente a la regla de selección determinada para coincidir. La unidad de gestión de mensajes 11 integra elementos de información predeterminados en la plantilla de generación de información seleccionada para generar una parte del mensaje que se va a agregar.

[0042] La unidad de generación de explicación 12 realiza el procesamiento de generación de explicación y el procesamiento de gestión de expresiones. En el procesamiento de generación de explicación, la unidad de generación de explicación 12 compara el mensaje de recepción almacenado en la unidad de gestión de mensajes 11 con la información transmitida ya almacenada en la unidad de generación de explicación 12, y detecta la presencia o ausencia de nueva información no transmitida. La información no transmitida corresponde a una diferencia de la información transmitida en el mensaje de recepción almacenado en la unidad de gestión de mensajes 11. La información transmitida corresponde a información pasada no transmitida utilizada para generar el texto explicativo convertido en los datos de descripción de audio que ya se han emitido. La unidad de generación de explicación 12 genera un texto explicativo que indica la situación de la competición representada por la nueva información no transmitida detectada. La unidad de generación de explicación 12 emite el texto explicativo generado a la unidad de síntesis de habla 13. La unidad de generación de explicación 12 puede emitir el texto explicativo generado al dispositivo transmisor 16 (Fig. 1) en paralelo con la salida a la unidad de síntesis de habla 13.

[0043] En el procesamiento de gestión de expresiones, la unidad de generación de explicación 12 monitorea el estado de salida (es decir, el estado de expresión) de los datos de descripción de audio de la unidad de síntesis de habla 13, y espera hasta que se complete la salida. Por ejemplo, la unidad de generación de explicación 12 determina que los datos de descripción de audio se emiten desde el momento en que se emite el texto explicativo hasta el momento en que se introduce una notificación final de salida (descrita más adelante) desde la unidad de síntesis de habla 13. La unidad de síntesis de habla 13 determina que los datos de descripción de audio no se emiten (sin expresión) hasta que el siguiente texto explicativo se emite después de que se introduce la notificación final de salida. Después de que finaliza la salida de los datos de descripción de audio, la unidad de generación de explicación 12 almacena la información no transmitida utilizada para generar el texto explicativo utilizado para la conversión a los datos de descripción de audio para los cuales finaliza la salida, como una adición de la información transmitida. Por tanto, se actualiza la información transmitida. El procesamiento de actualización de la información transmitida se puede realizar durante un tiempo de espera hasta que se complete el procedimiento de salida en la unidad de síntesis de habla 13. Posteriormente, la unidad de generación de explicación 12 realiza el procesamiento de generación de explicación nuevamente en la información transmitida después de la actualización y el último mensaje de recepción.

[0044] Al generar el texto explicativo, la unidad de generación de explicación 12 puede utilizar una plantilla explicativa almacenada de antemano en la unidad de generación de explicación 12. La plantilla explicativa son datos que tienen un campo de entrada para un elemento de información, que es una parte previamente designada del mensaje como un espacio en blanco, e indica un patrón de oración (en otras palabras, un texto perforado) de un texto explicativo para explicar el elemento de información. La unidad de generación de explicación 12 selecciona una plantilla explicativa correspondiente según una regla de selección predeterminada basada en el tipo de información no transmitida, de las plantillas explicativas almacenadas. La unidad de generación de explicación 12 genera el texto explicativo mediante la integración de elementos de información predeterminados en la plantilla explicativa seleccionada.

65

[0045] La unidad de síntesis de habla 13 realiza el procesamiento de síntesis de habla en la entrada de texto explicativo de la unidad de generación de explicación 12 para generar datos de descripción de audio que indican el audio de expresión que representa el contenido del habla en el texto explicativo. La unidad de síntesis de habla 13 puede usar, por ejemplo, un procedimiento de síntesis de habla basado en corpus como un procedimiento de síntesis de habla conocido. Cuando se utiliza el procedimiento de síntesis de habla basado en corpus, la unidad de síntesis de habla 13 se refiere a un corpus de audio almacenado de antemano. La unidad de síntesis de habla 13 emite los datos de descripción de audio generados al dispositivo transmisor 16 a una velocidad de habla predeterminada (por ejemplo, 300 a 400 mora/min en japonés). Cuando se finaliza la salida de los datos de descripción de audio, la unidad de síntesis de habla 13 emite una notificación de fin de salida que indica el final a la unidad de generación de explicación 12.

[0046] Incluso cuando el texto explicativo está escrito en japonés, la información del elemento de una parte del texto explicativo, por ejemplo, el nombre del jugador de la competición puede incluir una cadena de caracteres compuesta por alfabetos. En ese caso, la unidad de generación de explicación 12 puede realizar el procesamiento de estimación de lectura de la información del elemento, a partir de la cadena de caracteres o la información de nacionalidad del jugador además de la misma, y determinar la lectura de la cadena de caracteres de la información del elemento. La unidad de síntesis de habla 13 utiliza, por ejemplo, literatura no patentada: Miyazaki, Kumano, Imai, «translation of personal names using nationality information», 15th Information Science and Technology Forum Proceedings (FIT2016), E-018 n.º 2, p. 145-146, 23 de agosto de 2016, como procedimiento de procesamiento de estimación de lectura. (Mensaje)

[0047] A continuación, se describirá un ejemplo de un mensaje según la presente realización.

[0048] La Fig. 3 es un diagrama que ilustra un ejemplo de un mensaje que incluye datos del estado del juego. En la presente realización, los datos de competición formados basados en el ODF se pueden utilizar como un mensaje. ODF es una fuente de datos que indica un mensaje descrito mediante lenguaje de marcado extensible (XML), que es un tipo de lenguaje de marcado. Mediante el uso de etiquetas definidas en el lenguaje de marcado, se describen las relaciones de inclusión y las relaciones paralelas entre la información de configuración y sus tipos y atributos. Sin embargo, en los ejemplos mostrados en las Figuras 3 y 4, la notación de la etiqueta se omite o simplifica, y se representa información sustancial.

[0049] Al comienzo de la Fig. 3, el «estado del juego» se describe como el tipo de mensaje, y «final de fútbol masculino» se describe como el juego diana.

[0050] Además, «Brasil» se describe como el equipo 1 que es una de las cartas de batalla, y «Alemania» se describe como el equipo 2 que es el otro.

[0051] «3» y «4» se describen como los puntajes de los equipos 1 y 2, respectivamente.

[0052] Por lo tanto, el mensaje que se muestra en la Fig. 3 indica que, como estado del juego, la final de fútbol masculino es un partido entre la selección brasileña y la alemana, y el puntaje es de 3-4.

[0053] La Fig. 4 muestra un ejemplo de un mensaje que incluye información de eventos. Al comienzo de la Fig. 4, la «lista de eventos» se describe como el tipo de mensaje, y la «final de fútbol masculino» se describe como el juego diana.

[0054] Evento [1], evento [2],... y evento [8] indican la información del primer, segundo,... y octavo evento, respectivamente.

[0055] En el evento [8], «primera mitad 27 minutos» se describe como el tiempo. «Tiro libre (apuntando directamente a la portería)» se describe como el tipo de evento. Como resultado, se describe «éxito». «Equipo: Brasil, nombre del jugador NM» se describe como el actor.

[0056] Por lo tanto, el mensaje que se muestra en la Fig. 4 indica que como el octavo evento, el jugador NM que pertenece a la selección brasileña apunta directamente a la portería por un tiro libre en 27 minutos desde el inicio del conjunto de la primera mitad de la final de fútbol masculino.

(Gestión de mensajes)

[0057] A continuación, se describirá el procesamiento de gestión de mensajes según la presente realización.

[0058] La Fig. 5 es un diagrama de flujo que ilustra un primer ejemplo de procesamiento de gestión de mensajes según la presente realización. Este ejemplo es un tipo básico de procesamiento de gestión de mensajes.

[0059] (Etapa S11) La unidad de gestión de mensajes 11 recibe la última versión del mensaje de la fuente de

distribución en ese momento. A continuación, el procedimiento procede a la etapa S12.

[0060] (Etapa S12) La unidad de gestión de mensajes 11 acumula la última versión recibida del mensaje. Posteriormente, el procedimiento vuelve a la etapa S11.

5

[0061] La Fig. 6 es un diagrama de flujo que ilustra un segundo ejemplo de procesamiento de gestión de mensajes según la presente realización. Este ejemplo se caracteriza porque hay una pluralidad de tipos de mensaje y un mensaje se clasifica en cualquier tipo de los tipos.

10 **[0062]** (Etapa S21) La unidad de gestión de mensajes 11 recibe la última versión del mensaje de la fuente de distribución en ese momento. A continuación, el procedimiento procede a la etapa S22.

[0063] (Etapa S22) La unidad de gestión de mensajes 11 determina el tipo de la última versión del mensaje en ese momento. En la determinación de tipo, la unidad de gestión de mensajes 11 puede hacer referencia a una etiqueta unida al mensaje o un parámetro establecido. A continuación, el procedimiento procede a la etapa S23.

15 **[0064]** (Etapa S23) La unidad de gestión de mensajes 11 acumula la última versión recibida del mensaje para cada tipo determinado. Posteriormente, el procedimiento vuelve a la etapa S21.

20 **[0065]** En la etapa S12 (Fig. 5) o la etapa S23 (Fig. 6), la unidad de gestión de mensajes 11 puede rechazar el mensaje acumulado en el pasado y reemplazarlo completamente con la última versión del mensaje.

[0066] La Fig. 7 es un diagrama de flujo que ilustra un tercer ejemplo de procesamiento de gestión de mensajes según la presente realización. Este ejemplo se caracteriza porque la información de evento correspondiente a un patrón se genera cuando el patrón de la parte variada del mensaje coincide con un patrón predeterminado.

25 **[0067]** (Etapa S31) La unidad de gestión de mensajes 11 recibe la última versión del mensaje de la fuente de distribución en ese momento. A continuación, el procedimiento procede a la etapa S32.

30 **[0068]** (Etapa S32) La unidad de gestión de mensajes 11 acumula la última versión recibida del mensaje. En este momento, la unidad de gestión de mensajes 11 mantiene el mensaje acumulado en el pasado sin rechazarlo. A continuación, el procedimiento procede a la etapa S33.

[0069] (Etapa S33) La unidad de gestión de mensajes 11 compara la última versión del mensaje con la versión inmediatamente anterior del mensaje, y detecta una parte variada de la versión inmediatamente anterior. A continuación, el procedimiento procede a la etapa S34.

40 **[0070]** (Etapa S34) La unidad de gestión de mensajes 11 determina si hay o no un patrón que coincida con el patrón de parte variada detectado entre los patrones de parte variada almacenados de antemano. La unidad de gestión de mensajes 11 almacena un patrón de parte variada y un conjunto de plantilla de generación de información correspondiente al patrón para cada tipo de parte variada.

[0071] Cuando se determina que hay un patrón coincidente (SÍ en la etapa S34), el procedimiento procede a la etapa S35. Cuando se determina que no hay un patrón de coincidencia (NO en la etapa S34), el procedimiento

45

vuelve a la etapa S31.
[0072] (Etapa S35) La unidad de gestión de mensajes 11 especifica una regla de generación correspondiente al patrón de coincidencia con referencia a los datos de generación de eventos, y genera información de eventos en función de la regla de generación especificada y la última versión del mensaje. A continuación, el procedimiento

50

procede a la etapa S36.
[0073] (Etapa S36) La unidad de gestión de mensajes 11 agrega la información de evento generada a la última versión del mensaje, y la almacena en asociación con la última versión del mensaje. Posteriormente, el procedimiento

55

vuelve a la etapa S31.
(Procesamiento de generación de información de eventos)

[0074] A continuación, se describirá el procesamiento de generación de información de eventos realizado desde la etapa S33 hasta la etapa S35 (Fig. 7).

60

[0075] La Fig. 8 es un diagrama explicativo que muestra un ejemplo de procesamiento de generación de información de eventos.

[0076] La Fig. 8 muestra ejemplos de mensajes pasados, mensajes más recientes e información de eventos generados en este orden de arriba a abajo. En el ejemplo que se muestra en la Fig. 8, se asume que el mensaje de la

versión inmediatamente anterior y el mensaje de la última versión son datos del estado del juego en una competición de tenis, respectivamente.

- 5 **[0077]** En la etapa S33, la unidad de gestión de mensajes 11 compara la versión inmediatamente anterior del mensaje con la última versión del mensaje y detecta una parte variada. En el ejemplo que se muestra en la Fig. 8, los elementos de un recuento de juegos, la cantidad de errores no forzados y la cantidad de roturas de servicio (éxito de rotura de servicio) se detectan como partes variadas. Aquí, en el cuarto set (4S), se detecta el cambio en el recuento de juegos del competidor MR de 4 a 5, el cambio en el número de errores no forzados del competidor DP de 11 a 12, y el cambio en el número de roturas de servicio del competidor MR de 2 a 3.
- 10 **[0078]** En la etapa S34, la unidad de gestión de mensajes 11 detecta que hay un patrón cuyo tipo es un punto (puntuación) entre los patrones de la parte variada almacenada de antemano, como un patrón que coincide con el patrón de información de elemento de la parte variada detectada.
- 15 **[0079]** En la etapa S35, la unidad de gestión de mensajes 11 especifica una regla de generación correspondiente al patrón seleccionado con referencia a la plantilla de generación de información, y genera información de evento en función de la regla de generación especificada. Aquí, la unidad de gestión de mensajes 11 especifica al competidor MR como un actor incluido en la parte variada, el error no forzado del competidor DP que es el oponente del actor como causa, y el éxito de rotura del servicio del competidor MR (Y). A continuación, la unidad de gestión de mensajes 11 integra la información especificada con el identificador «[evento 1123]» para generar información de evento que incluye el tipo «punto», el actor «MR», la causa «error no forzado de DP» y «rotura de servicio: Y». El identificador [evento 1123] es información de identificación para identificar de forma única cada evento.
- 20 (Procesamiento de generación de explicación)
- 25 **[0080]** A continuación, se describirá el procesamiento de generación de explicación según la presente realización.
- 30 **[0081]** La Fig. 9 es un diagrama de flujo que ilustra un ejemplo de procesamiento de generación de explicación según la presente realización. El procedimiento que se muestra en la Fig. 9 se ejecuta de forma asíncrona con el procedimiento que se muestra en cualquiera de las Figs. 5 a 7.
- [0082]** (Etapa S41) La unidad de generación de explicación 12 compara el último mensaje almacenado en la unidad de gestión de mensajes 11 con la información transmitida almacenada en la propia unidad, y detecta información no transmitida. La unidad de generación de explicación 12 determina si existe o no un patrón que coincida con la información no transmitida detectada, entre los patrones de información no transmitida almacenados de antemano. La unidad de generación de explicación 12 almacena un conjunto de un patrón de información no transmitida y una plantilla explicativa correspondiente al patrón, para cada tipo de la información no transmitida. Cuando se determina que hay un patrón coincidente (SÍ en la etapa S41), el procedimiento procede a la etapa S43. Cuando se determina que no hay un patrón de coincidencia (NO en la etapa S41), el procedimiento procede a la etapa S42.
- 40 **[0083]** (Etapa S42) La unidad de generación de explicación 12 espera (duerme) sin realizar un procedimiento durante un tiempo determinado (por ejemplo, 1 segundo), y a continuación regresa al procedimiento de la etapa S41.
- 45 **[0084]** (Etapa S43) La unidad de generación de explicación 12 lee la plantilla explicativa correspondiente al patrón determinado para coincidir, y genera el texto explicativo en función de la plantilla explicativa leída y la información no transmitida. A continuación, el procedimiento procede a la etapa S44.
- 50 **[0085]** (Etapa S44) La unidad de generación de explicación 12 determina si el texto explicativo incluye o no un nombre de jugador descrito, por ejemplo, por una cadena de caracteres de alfabetos. Cuando se determina que se incluye el nombre del jugador (SÍ en la etapa S44), el procedimiento procede a la etapa S45. Cuando se determina que no se incluye un nombre de jugador (NO en la etapa S44), el procedimiento procede a la etapa S46.
- 55 **[0086]** (Etapa S45) La unidad de generación de explicación 12 realiza el procesamiento de estimación de lectura descrito anteriormente en función de la información de nacionalidad para la porción de nombre del jugador para determinar la lectura japonesa del texto explicativo. La unidad de generación de explicación 12 actualiza el texto explicativo con la lectura determinada para la parte. A continuación, el procedimiento procede a la etapa S46.
- 60 **[0087]** (Etapa S46) La unidad de generación de explicación 12 emite el texto explicativo generado a la unidad de síntesis de habla 13. La unidad de síntesis de habla 13 realiza el procesamiento de síntesis de habla en la entrada de texto explicativo de la unidad de generación de explicación 12 para generar datos de descripción de audio. La unidad de síntesis de habla 13 emite los datos de descripción de audio generados al dispositivo transmisor 16 a una velocidad de habla predeterminada. La unidad de generación de explicación 12 espera hasta que se complete la salida de los datos de descripción de audio, y a continuación procede al procedimiento de la etapa S47.
- 65

[0088] (Etapa S47) La unidad de generación de explicación 12 actualiza la información transmitida, acumulando la información no transmitida utilizada para generar el texto explicativo de salida como nueva información transmitida. Posteriormente, el procedimiento vuelve a la etapa S41.

5

[0089] El patrón de la información no transmitida puede incluir información designada que indica que se considera que un elemento de información predeterminado de la información transmitida no se transmite después de un tiempo predeterminado (por ejemplo, 30 segundos) ha transcurrido. En el procedimiento de la etapa S41, la unidad de generación de explicación 12 incluye el elemento de información designado en la información no transmitida cuando se determina si el patrón que incluye la información designada coincide con el patrón de la información no transmitida. Por lo tanto, incluso cuando el elemento de información designado se transmite realmente, se puede transmitir de nuevo. Por lo tanto, incluso cuando no hay cambios en la situación del juego o el cambio es pequeño, es posible transmitir repetidamente el elemento de información designado, por ejemplo, el estado del juego. Por lo tanto, en tal caso, es posible evitar que el oyente olvide la situación de la competición. Incluso un oyente que ha escuchado la descripción de audio desde la mitad del juego puede entender el estado del juego.

10

15

[0090] Entre los procedimientos mostrados en la Fig. 9, los procedimientos de la etapa S44 y la etapa S45 no son esenciales y pueden omitirse.

20 (Plantilla explicativa)

[0091] A continuación, se describirá un ejemplo de una plantilla explicativa según la presente realización.

[0092] La Fig. 10 es un diagrama que muestra un ejemplo de una plantilla explicativa según la presente realización.

25

[0093] La Fig. 10 es un ejemplo de una plantilla explicativa asociada con un patrón que tiene el tipo de evento como un punto. Esta plantilla explicativa muestra un patrón de oración para explicar un punto causado por la acción del jugador B como un evento en la competición y un recuento de juegos, el número de juegos adquiridos y el número de roturas de servicio en el set diana como un estado de juego. La plantilla explicativa proporcionada con, entre los elementos de información de la información no transmitida, espacios en blanco para sustituir a [nombre completo del jugador B] que obtiene una puntuación, [causa], [nombre completo del jugador B] que tuvo éxito en la rotura del servicio del juego que es la causa de la puntuación, [número total de juegos], [apellido del jugador A] y [apellido del jugador B] que son competidores, [número de juegos adquiridos por el jugador A] y [número de juegos adquiridos por el jugador B] como puntuación de cada competidor, y [número de roturas del servicio del jugador B] que es un actor en el momento de la puntuación.

30

35

[0094] A continuación, se describirá un ejemplo de una regla de selección de patrones correspondiente a la plantilla explicativa que se muestra en la Fig. 10.

40

[0095] La Fig. 11 es un diagrama que muestra un ejemplo de una regla de selección de plantilla explicativa que se muestra en la Fig. 10.

[0096] En este ejemplo, la condición para seleccionar el patrón correspondiente a la plantilla explicativa incluye, como el elemento de información incluido en la información no transmitida, el tipo de evento es «punto», y «error no forzado del jugador A (o jugador B)» y «éxito de rotura de servicio (Y)» como la causa del evento. Dicha regla de selección se establece de antemano para cada patrón de cada plantilla explicativa.

45

(Texto explicativo)

50

[0097] A continuación, se describirá un ejemplo de texto explicativo generado utilizando la plantilla explicativa que se muestra en la Fig. 10. La Fig. 12 es un diagrama que ilustra un ejemplo de un texto explicativo generado utilizando la plantilla explicativa que se muestra en la Fig. 10.

[0098] En la etapa S43 (Fig. 9), la unidad de generación de explicación 12 especifica la regla de selección del patrón mostrado en la Fig. 11 entre los patrones almacenados de antemano, como un patrón que coincide con el elemento de información incluido en la información no transmitida, como el patrón de información no transmitida. La unidad de generación de explicación 12 selecciona la plantilla explicativa ilustrada en la Fig. 12 como la plantilla explicativa correspondiente al patrón especificado.

55

60

[0099] La unidad de generación de explicación 12 sustituye el elemento de información correspondiente de la información no transmitida en cada uno de los espacios en blanco proporcionados en la plantilla explicativa seleccionada. Específicamente, la unidad de generación de explicación 12 sustituye los espacios en blanco de «nombre completo del jugador B», «causa», «apellido del jugador B» y «número total de juegos» en la plantilla explicativa que se muestra en la Fig. 10 con «AD·MR», «error no forzado de DP», «MR» y «10». La unidad de

65

generación de explicación 12 sustituye además los espacios en blanco de [apellido del jugador A], [apellido del jugador B], [número de juegos adquiridos por el jugador A], [número de juegos adquiridos por el jugador B] y [número de roturas de servicio del jugador B] con «DR», «MR», «5», «5» y «3». Por lo tanto, un texto explicativo que indica la oración «MR, éxito en la rotura de servicio del décimo juego. recuento de juegos, DR vs. MR, 5-5. Esta tercera rotura de servicio.» se genera.

[0100] Incluso cuando la información designada anteriormente se asocia con un determinado patrón, la unidad de generación de explicación 12 considera el elemento de información designado por la información de designación como información no transmitida después de que haya transcurrido un tiempo predeterminado desde la transmisión, y puede generar texto explicativo mediante el procedimiento descrito anteriormente.

[0101] Por ejemplo, se asume que la información del elemento que tiene un tipo de mensaje como estado de juego se establece en la información designada. En este caso, la información del elemento que tiene un tipo de mensaje como estado de juego se incluye en la información no transmitida, y se selecciona una plantilla explicativa correspondiente a un patrón que coincide con el patrón de información del elemento incluido en toda la información no transmitida. Aquí, como regla de selección, la unidad de generación de explicación 12 selecciona la siguiente plantilla explicativa como la plantilla explicativa correspondiente, cuando un tipo de estado de juego es «recuento de juegos», y la información del elemento coincide con las condiciones, incluyendo todas las de [nombre de juego], [apellido del jugador A], [apellido del jugador B], [número de sets], [número de juegos adquiridos por el jugador A] y [número de juegos adquiridos por el jugador B].

[0102] [nombre de la competición], [apellido del jugador A] vs. [apellido del jugador B]. actualmente el [número de sets]-ésimo. recuento de juegos [número de juegos adquiridos por el jugador A] vs. [número de juegos adquiridos por el jugador B].

[0103] La unidad de generación de explicación 12 sustituye los campos respectivos de [nombre del juego], [apellido del jugador A], [apellido del jugador B], [número de sets], [número de juegos adquiridos por el jugador A] y [número de juegos adquiridos por el jugador B] con «final individual de tenis para hombres», «DP», «MR», «1», «5» y «5», que es la información del elemento incluida en la información no transmitida, y cada uno de los cuales es la información del elemento incluida en la información no transmitida. Por consiguiente, la unidad de generación de explicación 12 puede generar el siguiente texto explicativo.

[0104] «Tenis individual masculino final, DP vs. MR. Actualmente el primer set. Cuenta de competición 5 a 5.»

[0105] Tal como se describió anteriormente, el dispositivo de generación de orientación de audio 10 según la presente realización incluye una unidad de gestión de mensajes 11, una unidad de generación de explicación 12 y una unidad de síntesis de habla 13. La unidad de gestión de mensajes 11 acumula los datos de competición que indican la situación de una competición como un mensaje de recepción. La situación de la competición descrita en los datos de la competición podrá estar representada en una difusión de vídeo en paralelo. La unidad de generación de explicación 12 genera un texto explicativo de la información no transmitida detectada a partir del mensaje de recepción, en función de la información transmitida. La unidad de síntesis de habla 13 emite audio convertido del texto explicativo. La unidad de generación de explicación 12 puede generar un nuevo texto explicativo de la información no transmitida, utilizando la información no transmitida utilizada para generar el texto explicativo como la información transmitida, después de que se completa la salida del audio de la unidad de síntesis de habla 13.

[0106] Según esta estructura, el audio basado en el texto explicativo que indica la situación de la competición representado en el vídeo se proporciona sin intervención humana. Además, una vez finalizado el suministro único de audio sin interrupción, se inicia en ese momento el suministro de audio para explicar la situación de una nueva competición. Por lo tanto, la situación de la competición se puede transmitir con precisión al oyente junto con el progreso de la competición que se muestra en el vídeo.

[0107] Cuando la información del evento que indica el evento de la competición no se incluye en la parte variada del mensaje de recepción, la unidad de gestión de mensajes 11 puede generar información del evento en la que la variación en el estado del juego es el estado de la competición incluida en la parte variada como un evento.

[0108] Según esta configuración, incluso cuando no se proporciona la información del evento, el estado de juego cambiado se puede transmitir al oyente, con el cambio en el estado del juego como un evento. Dado que el oyente puede entender el progreso de la competición con más detalle, la competición se puede disfrutar más.

[0109] La unidad de generación de explicación 12 puede manejar, entre piezas de información mantenidas como la información transmitida, la información para la cual ha transcurrido un tiempo predeterminado después de que se transmite, como no transmitida incluso si no hay cambios con respecto al último mensaje, y generar un texto que explique nuevamente el mismo contenido.

[0110] Según esta configuración, en un caso en el que la situación de la competición no cambia con el tiempo,

se puede transmitir repetidamente información que el oyente puede haber olvidado. Incluso un oyente que comienza a escuchar desde la mitad de la competición puede entender la situación de la competición.

5 **[0111]** La unidad de generación de explicación 12 puede almacenar de antemano una plantilla explicativa que indica el patrón de oración del texto explicativo para cada tipo de información no transmitida que se va a transmitir, y seleccionar una plantilla explicativa correspondiente al tipo de información no transmitida de la plantilla explicativa almacenada. La unidad de generación de explicación 12 genera el texto explicativo mediante la integración de la información no transmitida en la plantilla seleccionada.

10 **[0112]** Según esta configuración, sin desarrollar un sistema que soporte diferentes competiciones, solo se describe previamente la plantilla explicativa para cada competición, por lo que es posible transmitir la situación de la competición sin realizar un procesamiento complicado. La realización económica se puede lograr reduciendo el trabajo y la escala del sistema para ese propósito.

15 **[0113]** El sistema de difusión 1 también puede incluir el dispositivo de generación de orientación de audio 10 y un dispositivo transmisor 16 que transmite el texto explicativo y el audio generado.

20 **[0114]** Según esta configuración, la situación de la competición se transmite al espectador a través del texto explicativo que indica el contenido del audio a transmitir. Por lo tanto, incluso una persona con discapacidad auditiva puede ver el texto explicativo para comprender la situación de la competición. Por lo tanto, es posible hacer que los espectadores económicamente diversificados disfruten de la competición deportiva. <Segunda realización

25 **[0115]** A continuación, se describirá una segunda realización de la presente invención. El dispositivo de generación de orientación de audio 10 según la primera realización genera un comentario de audio (en otras palabras, descripción de audio) a partir de datos de competición que indican el estado de progreso tal como puntuaciones, goles e infracciones de competición deportiva. Al generar una descripción de audio a partir de estos datos de la competición, por ejemplo, es posible que las personas con discapacidad visual disfruten más de las difusiones deportivas en la televisión, y agregar comentarios de audio a los vídeos de competición deportiva transmitidos en vivo a través de Internet.

30 **[0116]** En los últimos años, la tecnología de síntesis de habla que utiliza modelos estadísticos ha avanzado. Por ejemplo, se ha propuesto una técnica en la que se utiliza una red neuronal profunda (DNN) para aprender las características de los hablantes y las expresiones emocionales, y a partir de un texto, se realiza una síntesis de habla en la que se controlan los hablantes y las emociones. Esto permite sintetizar el habla con una expresión rica.

35 **[0117]** Como se describe en la primera realización, el progreso de la competición puede guiarse por el audio mediante el uso de un procedimiento para generar una descripción de audio a partir de los datos de la competición. Sin embargo, en la primera realización, el texto es simplemente sintetizado por habla. Por lo tanto, el habla sintetizada se convierte en audio sin entonación nítida, y la intención de la orientación puede no expresarse lo suficiente.

40 **[0118]** Al generar una audiodescripción utilizando un procedimiento que sintetiza el habla en el que se controlan el hablante y las emociones, simplemente controlar al hablante y las emociones para todo el texto no se puede controlar el audio y la entonación acentuada en el medio del texto. Por lo tanto, incluso cuando se genera una descripción de audio a partir de los datos de la competición utilizando este procedimiento, existe un problema de que 45 la intención de la orientación no puede expresarse lo suficiente porque no es posible agregar una nitidez a la entonación del audio.

50 **[0119]** La presente realización se ha realizado en vista de tal problema, y un objetivo de la misma es generar audio con entonación a partir de datos de competición. En la siguiente descripción, se describirá principalmente una diferencia con respecto a la primera realización. Los mismos elementos constituyentes que en la primera realización se indican con los mismos números de referencia, y se utiliza la descripción.

55 **[0120]** La Fig. 13 es un diagrama de bloques que ilustra un ejemplo de configuración de un dispositivo de generación de orientación de audio 30 según la presente realización.

[0121] El dispositivo de generación de orientación de audio 30 genera audio (voz sintetizada) a partir de datos de competición que indican el estado de progreso de los deportes o similares.

60 **[0122]** En la etapa previa al aprendizaje, el dispositivo de generación de orientación de audio 30 utiliza un corpus como datos de aprendizaje almacenados en el dispositivo de almacenamiento de corpus 40 para aprender un modelo de audio para convertir texto en habla sin conexión mediante DNN.

[0123] El corpus incluye un corpus de audio de entonación acentuada y un corpus de audio normal.

65 **[0124]** El corpus de audio de entonación acentuada es el aprendizaje de datos en los que un texto y un habla

hablada con entonación acentuada en el texto (audio de entonación acentuada) se emparejan.

[0125] El corpus de audio normal es el aprendizaje de datos en los que un texto y un habla hablada normalmente (sin entonación acentuada) en el texto (audio normal) se emparejan.

5

[0126] El dispositivo de generación de orientación de audio 30 genera audio explicando el progreso de la competición en línea, a partir de los datos de la competición, utilizando el modelo de audio aprendido, en la etapa de procesamiento de síntesis.

10 **[0127]** El dispositivo de generación de orientación de audio 30 genera un texto que explica el progreso de la competición con un estilo de habla que indica la porción acentuada (texto con el estilo de habla) a partir de los datos de competición, en función de una plantilla en la que se determina de antemano la parte a acentuar de los datos de competición, y realiza la síntesis de habla utilizando el modelo de audio aprendido.

15 **[0128]** Por lo tanto, el dispositivo de generación de orientación de audio 30 puede generar audio con una entonación nítida, a partir de los datos de la competición.

[0129] El dispositivo de generación de orientación de audio 30 incluye una unidad de gestión de mensajes 11, una unidad de generación de explicación 12, una unidad de preaprendizaje 31 y una unidad de síntesis de habla 34.

20

[0130] La unidad de preaprendizaje 31 aprende varios parámetros que constituyen el modelo de audio en la etapa de preaprendizaje.

25 **[0131]** La unidad de síntesis de habla 34 genera audio (habla sintetizada) a partir de los datos de la competición, utilizando el modelo de audio aprendido en la etapa de procesamiento de síntesis.

[0132] El dispositivo de generación de orientación de audio 30 está conectado al dispositivo de almacenamiento de corpus 40 de modo que se pueden transmitir y recibir diversos datos.

30 (Unidad de preaprendizaje)

[0133] A continuación, se describirá la configuración de la unidad de preaprendizaje 31 con referencia a la Fig. 14. La unidad de preaprendizaje 31 incluye una unidad de entrada de corpus 32 y una unidad de aprendizaje modelo 33.

35

[0134] La unidad de entrada de corpus 32 introduce un corpus de audio de entonación acentuada y un corpus de audio normal desde el dispositivo de almacenamiento de corpus 40.

40 **[0135]** La unidad de entrada de corpus 32 lee texto y audio (en otras palabras, audio de entonación acentuada o audio normal) en asociación entre sí desde el dispositivo de almacenamiento de corpus 40, y emite el texto y audio a la unidad de aprendizaje modelo 33. La unidad de entrada de corpus 32 emite la información de discriminación de entonación acentuada que indica que el audio correspondiente al texto leído del corpus de audio de entonación acentuada es audio de entonación acentuada en asociación con el texto y el audio a la unidad de aprendizaje modelo 33. La unidad de entrada de corpus 32 emite la información de discriminación de entonación acentuada que indica
45 que el audio correspondiente al texto leído del corpus de audio normal es audio normal (no es audio de entonación acentuada) en asociación con el texto y el audio a la unidad de aprendizaje modelo 33.

50 **[0136]** La unidad de aprendizaje modelo 33 aprende el modelo de duración del tiempo 346a y el modelo de rasgo acústico 346b, como el modelo de audio DNN, usando la entrada de texto de la unidad de entrada de corpus 32, información de discriminación de entonación acentuada y audio (en otras palabras, audio de entonación acentuada o audio normal). La unidad de aprendizaje modelo 33 incluye una unidad de análisis de texto 331, una unidad de análisis de audio 332 y una unidad de aprendizaje de DNN 333.

55 **[0137]** La unidad de análisis de texto 331 analiza el texto para generar un rasgo de lenguaje para cada fonema y trama, y genera etiquetas (información de entonación acentuada e información de terminación de oración) para cada fonema que especifica el estilo de habla, en función de la información de discriminación de entonación acentuada. La unidad de análisis de texto 331 incluye una unidad de generación de rasgo de lenguaje de fonema 331a y una unidad de generación de rasgo de lenguaje de trama 331b.

60 **[0138]** La unidad de generación de rasgo de lenguaje de fonema 331a genera un rasgo de lenguaje de fonema que es un rasgo de lenguaje para cada uno de los fonemas que constituyen el texto de entrada. El rasgo de lenguaje de fonema indica un rasgo dependiendo del contexto del fonema. Por ejemplo, el rasgo de lenguaje de fonema es uno o un conjunto de información de fonemas (etiqueta de fonema que incluye fonemas anteriores y siguientes), posición de acento, información de frase de acento, parte de información de habla, información de párrafo de exhalación y
65 similares. Dado que estos rasgos de lenguaje de fonema se pueden obtener mediante análisis de texto general

conocido públicamente (por ejemplo, análisis morfológico, estimación de prosodia o similares), se omite la descripción detallada del procedimiento de análisis.

5 **[0139]** La unidad de generación de rasgo de lenguaje de fonema 331a agrega etiquetas específicas de estilo de habla (etiqueta de estilo de habla [información de entonación acentuada e información de terminación de oración]) que indica la entonación acentuada y termina una oración con una frase nominal al rasgo de lenguaje de fonema generado.

10 **[0140]** En un caso en el que la información de discriminación de entonación acentuada que se introduce junto con el texto es información que indica que el audio es audio de entonación acentuada, la unidad de generación de rasgo de lenguaje de fonema 331a agrega, como etiqueta de estilo de habla, una etiqueta que indica la entonación acentuada (información de entonación acentuada; con entonación acentuada). En un caso en el que la información de discriminación de entonación acentuada que se introduce junto con el texto es información que indica que el audio es audio normal (en otras palabras, no es audio de entonación acentuada), la unidad de generación de rasgo de lenguaje de fonema 331a agrega, como una etiqueta de estilo de habla, una etiqueta que indica normal (información de entonación acentuada; sin entonación acentuada) a todos los fonemas que constituyen el texto.

20 **[0141]** En un caso en el que la parte del habla de la palabra de terminación de oración del texto de entrada es un sustantivo, la unidad de generación de rasgo de lenguaje de fonema 331a agrega, como una etiqueta de estilo del habla, una etiqueta que indica la terminación de una oración con una frase nominal (información de terminación de oración; con la terminación de una oración con una frase nominal) a todos los fonemas que constituyen la palabra de terminación de la oración. La unidad de generación de rasgo de lenguaje de fonema 331a agrega una etiqueta que indica que no es la terminación de una oración con una frase nominal (información de terminación de oración; sin terminar una oración con una frase nominal) a fonemas distintos de los fonemas a los que una etiqueta indica la terminación de una oración con una frase nominal. La terminación de una oración con una frase nominal significa una forma de expresión en la que una parte del habla de una palabra al final de una oración es un sustantivo en una oración japonesa. En las oraciones japonesas, un predicado generalmente se coloca al final de la oración. Como predicado, se utilizan palabras cuya parte del habla es un verbo, un adjetivo, un verbo adjetivo o un verbo auxiliar. En otras palabras, una oración terminada con una frase nominal puede considerarse como una forma de expresión en la que se omite el predicado que representa la acción o el estado del sustantivo al final de la oración. Por lo tanto, la entonación puede ser diferente de la oración en la que la palabra de terminación de oración es un predicado. En general, la frecuencia de la terminación de una oración con una frase nominal es rara en las declaraciones japonesas. Sin embargo, en la declaración utilizada en la situación de deportes en vivo y similares, terminar una oración con una frase nominal tiende a usarse con relativa frecuencia. La unidad de preaprendizaje 31 según la presente realización presta atención a este hecho y utiliza la información que indica la presencia o ausencia de la terminación de una oración con una frase nominal para el aprendizaje del modelo de audio.

40 **[0142]** La unidad de generación de rasgo de lenguaje de fonema 331a emite el rasgo de lenguaje de fonema y las etiquetas de estilo de habla (información de entonación acentuada e información de terminación de oración) a la unidad de generación de rasgo de lenguaje de trama 331b y la unidad de aprendizaje de modelo de duración del tiempo 333a de la unidad de aprendizaje de DNN 333.

45 **[0143]** La unidad de generación de rasgo de lenguaje de fonema 331a emite la etiqueta de fonema específica de fonema analizada al separador de fonemas 332a de la unidad de análisis de audio 332 para cada fonema.

50 **[0144]** La unidad de generación de rasgo de lenguaje de trama 331b genera un rasgo de lenguaje de trama que es un rasgo de lenguaje correspondiente a una trama que constituye una sección de duración del tiempo del fonema del texto de entrada. Una trama es una unidad para generar un rasgo acústico analizado por la unidad de análisis de audio 332.

55 **[0145]** El rasgo de lenguaje de trama se forma al incluir el rasgo temporal del fonema, tal como los números de tramas que constituyen la sección de duración del tiempo del fonema (es decir, la posición de la trama) y la cantidad total de tramas en el rasgo de lenguaje de fonema generado por la unidad de generación de rasgo de lenguaje de fonema 331a. La unidad de generación de rasgo de lenguaje de trama 331b genera el rasgo de lenguaje de trama correspondiente a la duración del tiempo de fonema para un fonema.

60 **[0146]** La unidad de generación de rasgo de lenguaje de trama 331b recibe la duración del tiempo de fonemas de la unidad de análisis de audio 332 y obtiene la cantidad de tramas que constituyen la sección de la duración del tiempo de fonema, la cantidad total de tramas y similares. La duración del fonema es, por ejemplo, el número de tramas por fonema.

65 **[0147]** La unidad de generación de rasgo de lenguaje de trama 331b emite los rasgos de lenguaje de trama generados para la cantidad de tramas correspondientes a un fonema y la entrada de etiqueta de estilo de habla de la unidad de generación de rasgo de lenguaje de fonema 331a en asociación entre sí a la unidad de aprendizaje de modelo de rasgo acústico 333b de la unidad de aprendizaje DNN 333.

- [0148]** La unidad de análisis de audio 332 analiza el audio de entrada y genera una duración del tiempo de fonema y un rasgo acústico para cada trama. La unidad de análisis de audio 332 incluye un separador de fonemas 332a y una unidad de generación de rasgo acústico 332b.
- 5 **[0149]** El separador de fonemas 332a detecta una ruptura de fonemas de la entrada de etiqueta de fonemas de la unidad de análisis de texto 331 en la forma de onda del habla de la entrada de audio a la unidad de entrada de corpus 32.
- 10 **[0150]** El separador de fonemas 332a detecta una posición de ruptura del fonema usando, por ejemplo, un diccionario de segmentos (no mostrado) en el que un fonema está asociado con una forma de onda del fonema.
- [0151]** El separador de fonemas 332a emite la posición de ruptura detectada del fonema a la unidad de generación de rasgo acústico 332b.
- 15 **[0152]** El separador de fonemas 332a obtiene la duración del tiempo del fonema correspondiente a la etiqueta de fonema, en función de la posición de ruptura detectada del fonema, y emite la duración del tiempo a la unidad de generación de rasgo de lenguaje de trama 331b de la unidad de análisis de texto 331, y la unidad de aprendizaje de modelo de duración del tiempo 333a de la unidad de aprendizaje de DNN 333.
- 20 **[0153]** La duración del tiempo del fonema generado por el separador de fonemas 332a son datos correctos cuando la unidad de aprendizaje del modelo de duración del tiempo 333a aprende el modelo de duración del tiempo 346a.
- 25 **[0154]** La unidad de generación de rasgo acústico 332b genera un rasgo acústico para cada trama del audio de entrada. La unidad de generación de rasgo acústico 332b genera rasgos acústicos para el número de tramas que constituyen la sección de la duración del tiempo del fonema. La unidad de generación de rasgo acústico 332b realiza análisis acústicos cortando el audio de entrada en una longitud de trama predeterminada (por ejemplo, 25 milisegundos) cada período de trama predeterminado (por ejemplo, 5 milisegundos) en la sección de la duración del tiempo del fonema en función de la posición de ruptura detectada por el separador de fonemas 332a, y genera, como rasgo acústico, coeficientes espectrales (tales como coeficientes de mel cepstrum), tono (tal como frecuencia de tono logarítmico), coeficiente de ruido, valor de determinación con/sin voz y similares.
- 30 **[0155]** Por ejemplo, la unidad de generación de rasgo acústico 332b genera un total de 62 características estáticas dimensionales que incluyen el coeficiente espectral dimensional 60, un tono unidimensional y un coeficiente de ruido unidimensional, 124 características dinámicas dimensionales de la diferencia de primer orden y la diferencia de segundo orden en la dirección temporal de las características estáticas, y un total de 187 rasgos acústicos dimensionales que consisten en valores de determinación unidimensionales con/sin voz. Dado que estos rasgos acústicos se pueden obtener mediante análisis acústico general, se omite la descripción detallada sobre un procedimiento de análisis.
- 40 **[0156]** La unidad de generación de rasgo acústico 332b emite el rasgo acústico generado a la unidad de aprendizaje de modelo de rasgo acústico 333b de la unidad de aprendizaje de DNN 333.
- 45 **[0157]** El rasgo acústico generado por la unidad de generación de rasgo acústico 332b son datos correctos cuando la unidad de aprendizaje de modelo de rasgo acústico 333b aprende el modelo de rasgo acústico 346b.
- [0158]** La unidad de aprendizaje de DNN 333 aprende un modelo de duración del tiempo 346a y un modelo de rasgo acústico 346b como modelos de audio de DNN. La unidad de aprendizaje de DNN 333 incluye una unidad de aprendizaje de modelo de duración del tiempo 333a y una unidad de aprendizaje de modelo de rasgo acústico 333b.
- 50 **[0159]** La unidad de aprendizaje del modelo de duración del tiempo 333a aprende el modelo de duración del tiempo 346a como un modelo DNN que recibe el rasgo de lenguaje de fonema y las etiquetas de estilo de habla (información de entonación acentuada e información de terminación de oración) y emite la duración del tiempo del fonema.
- 55 **[0160]** La Fig. 16 muestra un ejemplo del modelo de duración del tiempo 346a configurado por DNN. Como se muestra en la Fig. 16, el modelo de duración del tiempo 346a puede configurarse mediante una red neuronal de avance de alimentación (FFNN) que incluye una capa de entrada I, una capa oculta H y una capa de salida O.
- 60 **[0161]** La unidad de aprendizaje del modelo de duración del tiempo 333a ingresa a cada unidad de la capa de entrada I del modelo de duración del tiempo 346a, el rasgo de lenguaje de fonema analizado por la unidad de análisis de texto 331, y las etiquetas de la información de entonación acentuada (presencia o ausencia de entonación acentuada) y la información de terminación de oración (presencia o ausencia de terminación de una oración con una frase nominal) que son etiquetas de estilo de habla. La unidad de aprendizaje del modelo de duración del tiempo 333a
- 65

propaga rasgos de lenguaje de fonemas y valores de etiquetas de estilo de habla (información de entonación acentuada e información de terminación de oración) de entrada a la capa de entrada I con pesos, y aprende los parámetros para ser pesos en el modelo de duración del tiempo 346a de modo que la salida de la capa de salida O se convierte en la duración del tiempo del fonema analizado por la unidad de análisis de audio 332. Para el aprendizaje de parámetros, por ejemplo, se puede usar un procedimiento de propagación de retorno de error.

[0162] La unidad de aprendizaje del modelo de duración del tiempo 333a actualiza secuencialmente el modelo de duración del tiempo 346a almacenado en la unidad de almacenamiento de modelo 346 al modelo de duración del tiempo 346a.

10

[0163] La unidad de aprendizaje de modelo de rasgo acústico 333b aprende el modelo de rasgo acústico 346b como un modelo DNN que recibe el rasgo de lenguaje de trama y las etiquetas de estilo de habla (información de entonación acentuada e información de terminación de oración) y emite un rasgo acústico para cada trama de fonema.

15 **[0164]** La Fig. 17 muestra un ejemplo del modelo de rasgo acústico 346b configurado por DNN. Como se muestra en la Fig. 17, el modelo de rasgo acústico 346b puede configurarse mediante una red neuronal de propagación directa (FFNN) que incluye una capa de entrada I, una capa oculta H y una capa de salida O.

20 **[0165]** La unidad de aprendizaje de modelo de rasgo acústico 333b ingresa a cada unidad de la capa de entrada I del modelo de rasgo acústico 346b, el rasgo de lenguaje de trama analizado por la unidad de análisis de texto 331, y etiquetas de la información de entonación acentuada (presencia o ausencia de entonación acentuada) e información de terminación de oración (presencia o ausencia de terminación de una oración con una frase nominal) que son etiquetas de estilo de habla. La unidad de aprendizaje de modelo de rasgo acústico 333b propaga cada valor de rasgo de lenguaje de trama y etiquetas de estilo de habla (información de entonación acentuada e información de terminación de oración) de entrada a la capa de entrada I con pesos, y aprende los parámetros para ser pesos en el modelo de rasgo acústico 346b de modo que la salida de la capa de salida O se convierte en el rasgo acústico de la trama analizada por la unidad de análisis de audio 332. Para el aprendizaje de parámetros, el procedimiento de propagación inversa de errores se puede utilizar como en el modelo de duración del tiempo 346a.

30 **[0166]** La unidad de aprendizaje de modelo de rasgo acústico 333b actualiza secuencialmente el modelo de rasgo acústico 346b almacenado en la unidad de almacenamiento de modelo 346 al modelo de rasgo acústico 346b obtenido mediante aprendizaje.

35 **[0167]** La unidad de almacenamiento de modelo 346 almacena el modelo de audio (en otras palabras, el modelo de duración del tiempo 346a y el modelo de rasgo acústico 346b) aprendido por la unidad de aprendizaje de modelo 33. La unidad de almacenamiento de modelo 346 puede configurarse mediante un medio de almacenamiento general tal como una memoria semiconductor.

40 **[0168]** El modelo de audio (en otras palabras, el modelo de duración del tiempo 346a y el modelo de rasgo acústico 346b) aprendido en la etapa de aprendizaje previo, almacenado en la unidad de almacenamiento de modelo 346, es utilizado por la unidad de generación de habla sintetizada 341.

45 **[0169]** La unidad de preaprendizaje 31 descrita anteriormente puede generar un modelo de audio en el que la duración del tiempo del fonema y el rasgo acústico se aprenden en audio cuya entonación está acentuada que el audio normal, y audio de terminación de una oración con una frase nominal, además del audio normal.

[0170] A continuación, se describirá otro ejemplo de los datos de la competición con referencia a las Figs. 18 y 19.

50 **[0171]** Los datos de la competición son, por ejemplo, datos correspondientes a los jugadores que participan en la competición (Fig. 18), y el contenido de la competición y el progreso de la competición, como una lista de eventos que ocurren durante la competición (Fig. 19). Las Figs. 18 y 19 muestran ejemplos de datos de competición para tenis, como ejemplo.

55 **[0172]** La Fig. 18 muestra que [nombre completo del jugador A] es «A1·A2», [apellido del jugador A] es «A2» y similares.

60 **[0173]** Además, la Fig. 19 muestra que el evento 123 se produjo en [hora] O hora Δ minuto segundo. [Set] indica el número de sets de juego. Los datos de la competición se añaden secuencialmente en la unidad de gestión de mensajes 11 cada vez que se actualiza el evento. Aquí, el evento 123 ocurre durante el cuarto set («4S»). [Recuento de juego] indica que el recuento de juegos se ha actualizado. Aquí, indica que «B2» (jugador B) tiene el quinto punto («5») [Error no forzado] indica que ha ocurrido un error no forzado. Aquí, «A2» (jugador A) ha cometido un error no forzado, y el error no forzado es el «duodécimo» error. [Éxito de rotura de servicio] indica que la rotura de servicio ha sido exitosa. Aquí, «B2» (jugador B) ha tenido éxito en la «sexta» rotura de servicio.

65

[0174] Además, la Fig. 19 ilustra [número total de juegos], [número de juegos adquiridos por el jugador A], [número de juegos adquiridos por el jugador B], [número de roturas de servicio del jugador B] y similares como situación de las competiciones.

5 **[0175]** Por supuesto, los datos de competición ilustrados en las Figuras 18 y 19 no se limitan a esto, y los datos de competición pueden diferir según el tipo de competición.

[0176] La unidad de generación de explicación 12 almacena de antemano una plantilla explicativa para convertir los datos de la competición en texto para la síntesis de habla. Una unidad de almacenamiento de plantilla
10 (no mostrada) que almacena la plantilla explicativa puede configurarse mediante un medio de almacenamiento general tal como una memoria de semiconductor.

[0177] La unidad de almacenamiento de plantillas almacena una pluralidad de plantillas explicativas en asociación con condiciones que indican datos predeterminados en los datos de competición.
15

[0178] Por ejemplo, la Fig. 20 muestra una plantilla cuando el jugador A provoca un error no forzado y el jugador B tiene éxito en la rotura de servicio, en los datos de competición que se muestran en la Fig. 18.

[0179] En la plantilla explicativa que se muestra en la Fig. 20, la porción interpuesta entre «[«y »]» cambia según los datos de competición.
20

[0180] En esta plantilla explicativa, una porción en la que el habla se acentúa y se pronuncia se indica mediante una etiqueta predeterminada. Por ejemplo, como se muestra en la Fig. 20, una porción interpuesta entre la etiqueta <entonación acentuada> y la etiqueta </entonación acentuada> se establece como una diana de entonación
25 acentuada. Aquí, se muestra un ejemplo de acentuar «[número de roturas de servicio del jugador B] KAIME NO BREEKU».

[0181] En la plantilla explicativa, la presencia o ausencia de la entonación acentuada puede controlarse mediante la adición adicional de condiciones a las condiciones mostradas en la Fig. 20.
30

[0182] Por ejemplo, como se muestra en la Fig. 21, además del caso en el que el jugador A ha causado un error no forzado y el jugador B ha tenido éxito en la rotura de servicio mostrado en la Fig. 20), la plantilla explicativa utilizada en el caso en el que el número de roturas de servicio del jugador B es un número predeterminado (por ejemplo, «5») o más es la plantilla explicativa que acentúa «[número de roturas de servicio del jugador B] KAIME NO
35 BUREEKU» es entonación acentuada, como en la Fig. 20.

[0183] La plantilla explicativa utilizada en un caso donde el número de roturas de servicio del jugador B es menor que un número predeterminado (por ejemplo, «5») es, como se muestra en la Fig. 22, la plantilla explicativa obtenida al eliminar la etiqueta <entonación acentuada> y la etiqueta </entonación acentuada> de la plantilla
40 explicativa mostrada en la Fig. 21.

[0184] La unidad de generación de explicación 12 convierte la entrada de datos de competición a la unidad de gestión de mensajes 11 en un texto explicativo para realizar la síntesis de habla, utilizando la plantilla explicativa almacenada de antemano.
45

[0185] Por ejemplo, la unidad de generación de explicación 12 recibe los datos de la competición ilustrados en las Figs. 18 y 19, lee una plantilla explicativa en la que varios datos coinciden con las condiciones e incrusta una cadena de caracteres dependiendo de los datos de la competición. En caso de que se actualicen los datos de la competición, la unidad de generación de explicación 12 genera un texto explicativo a partir del contenido actualizado.
50

[0186] Por ejemplo, cuando se reciben los datos de la competición (aquí, el evento 123) que se muestra en la Fig. 19, la unidad de generación de explicación 12 incorpora una cadena de caracteres en la plantilla explicativa que se muestra en la Fig. 20 o la Fig. 21 para generar el texto explicativo que se muestra en la Fig. 23.

55 **[0187]** Cuando se utiliza la plantilla explicativa mostrada en la Fig. 22, la unidad de generación de explicación 12 genera el texto explicativo ilustrado en la Fig. 24.

[0188] Como resultado, la unidad de generación de explicación 12 puede generar, como el texto explicativo, el texto con el estilo de habla que indica la porción a acentuar en el texto como el estilo del habla con la etiqueta.
60

[0189] La unidad de generación de explicación 12 emite el texto generado con el estilo de habla a la unidad de síntesis de habla 13.
65

(Unidad de síntesis de habla)

[0190] A continuación, se describirá la configuración de la unidad de síntesis de habla 13. En el ejemplo ilustrado en la Fig. 13, la unidad de síntesis de habla 13 incluye una unidad de generación de habla sintetizada 341 y una unidad de almacenamiento modelo 346.

5 **[0191]** La unidad de generación de habla sintetizada 341 utiliza el modelo de audio aprendido almacenado en la unidad de almacenamiento de modelo 346 (en otras palabras, el modelo de duración del tiempo 346a y el modelo de rasgo acústico 346b) para realizar el procesamiento de síntesis de habla en el texto con el estilo de habla generado por la unidad de generación de explicación 12.

10 **[0192]** En el ejemplo ilustrado en la Fig. 15, la unidad de generación de habla sintetizada 341 incluye una unidad de análisis de estilo 342, una unidad de análisis de texto 343, una unidad de cálculo de DNN 344 y una unidad de síntesis 345.

[0193] La unidad de análisis de estilo 342 analiza la porción acentuada en el texto con el estilo de habla generado por la unidad de generación de explicación 12.

[0194] La unidad de análisis de estilo 342 especifica una porción acentuada en la oración mediante la detección de etiquetas (<entonación acentuada> y </entonación acentuada>) determinadas de antemano para cada oración en el texto con el estilo de habla. La unidad de análisis de estilo 342 emite en unidades de oraciones, el texto en el que se elimina la etiqueta del texto con el estilo de habla e información de posición de entonación acentuada que indica la porción acentuada en el texto, a la unidad de análisis de texto 343. Cuando no hay una etiqueta predeterminada en el texto con el estilo de habla, la unidad de análisis de estilo 342 emite, a la unidad de análisis de texto 343, el texto y la información de posición de entonación acentuada indicando que no hay una porción acentuada, en unidades de oraciones.

25 **[0195]** La unidad de análisis de texto 343 analiza el texto para generar un rasgo de lenguaje para cada fonema y trama, y genera una etiqueta para cada fonema que especifica el estilo de habla en función de la información de posición acentuada. La unidad de análisis de texto 343 incluye una unidad de generación de rasgo de lenguaje de fonema 343a y una unidad de generación de rasgo de lenguaje de trama 343b.

30 **[0196]** La unidad de generación de rasgo de lenguaje de fonema 343a genera un rasgo de lenguaje de fonema que es un rasgo de lenguaje para cada uno de los fonemas que constituyen el texto de entrada. El rasgo de lenguaje de fonema indica un rasgo de un fonema. La unidad de generación de rasgo de lenguaje de fonema 343a agrega etiquetas específicas de estilo de habla que indican la entonación acentuada y terminan una oración con una frase nominal (etiquetas de estilo de habla [información de entonación acentuada e información de terminación de oración]) al rasgo de lenguaje de fonema generado. La unidad de generación de rasgo de lenguaje de fonema 343a tiene la misma función que la unidad de generación de rasgo de lenguaje de fonema 331a descrita con referencia a la Fig. 14.

40 **[0197]** La unidad de generación de rasgo de lenguaje de fonema 331a añade, como etiqueta de estilo de habla, una etiqueta que indica la entonación acentuada (con entonación acentuada), a todos los fonemas que constituyen el texto correspondiente al audio de entonación acentuada. Por otro lado, la unidad de generación de rasgo de lenguaje de fonema 343a agrega, como una etiqueta de estilo de habla, una etiqueta que indica la entonación acentuada (información de entonación acentuada; con entonación acentuada), a los fonemas correspondientes a la porción acentuada indicada por la información de posición de entonación acentuada analizada por la unidad de análisis de estilo 342, y agrega una etiqueta que indica que no es la entonación acentuada (información de entonación acentuada; sin entonación acentuada) a los otros fonemas.

50 **[0198]** La unidad de generación de rasgo de lenguaje de fonema 343a emite el rasgo de lenguaje de fonema y las etiquetas de estilo de habla (la información de entonación acentuada y la información de terminación de oración) a la unidad de generación de rasgo de lenguaje de trama 343b y la unidad de cálculo de duración del tiempo 344a de la unidad de cálculo de DNN 344.

55 **[0199]** La unidad de generación de rasgo de lenguaje de trama 343b genera un rasgo de lenguaje de trama que es un rasgo de lenguaje correspondiente a un trama que constituye la sección de la duración del tiempo del fonema del texto de entrada. La unidad de generación de rasgo de lenguaje de trama 343b tiene la misma función que la unidad de generación de rasgo de lenguaje de trama 331b descrita con referencia a la Fig. 14.

60 **[0200]** La unidad de generación de rasgo de lenguaje de trama 343b recibe el rasgo de lenguaje de fonema generado por la unidad de generación de rasgo de lenguaje de fonema 343a y la duración del tiempo del fonema generado por la unidad de cálculo de duración del tiempo 344a de la unidad de cálculo DNN 344 para generar el rasgo de lenguaje de trama de la misma manera que la unidad de generación de rasgo de lenguaje de trama 331b (Fig. 14).

65 **[0201]** La unidad de generación de rasgo de lenguaje de trama 343b emite el rasgo de lenguaje de trama generado con la entrada de etiqueta de estilo de habla de la unidad de generación de rasgo de lenguaje de fonema 343a en asociación entre sí a la unidad de cálculo de rasgo acústico 344b de la unidad de cálculo DNN 344.

- 5 **[0202]** La unidad de cálculo de DNN 344 calcula DNN utilizando los modelos de audio (modelo de duración del tiempo 346a y modelo de rasgo acústico 346b) almacenados en la unidad de almacenamiento de modelo 346. La unidad de cálculo de DNN 344 incluye una unidad de cálculo de duración del tiempo 344a y una unidad de cálculo de rasgo acústico 344b.
- 10 **[0203]** La unidad de cálculo de duración del tiempo 344a genera la duración del tiempo del fonema a partir del rasgo de lenguaje de fonema y las etiquetas de estilo de habla (información de entonación acentuada e información de terminación de oración) usando el modelo de duración del tiempo aprendido 346a.
- 15 **[0204]** Por ejemplo, la unidad de cálculo de la duración del tiempo 344a ingresa el rasgo de lenguaje de fonema y las etiquetas de estilo de habla (la información de entonación acentuada y la información de terminación de la oración) a cada unidad de la capa de entrada I del modelo de duración del tiempo 346a que se muestra en la Fig. 16. La unidad de cálculo de duración del tiempo 344a calcula la DNN usando los parámetros del modelo de duración del tiempo aprendido 346a, y establece la salida de valor de la capa de salida O como la duración del tiempo del fonema.
- [0205]** La unidad de cálculo de duración del tiempo 344a emite la duración del tiempo de fonema generada a la unidad de generación de rasgo de lenguaje de trama 343b de la unidad de análisis de texto 343.
- 20 **[0206]** La unidad de cálculo de rasgo acústico 344b utiliza el modelo de rasgo acústico aprendido 346b para generar un rasgo acústico para cada trama de fonema, a partir del rasgo de lenguaje de trama y las etiquetas de estilo de habla (la información de entonación acentuada y la información de terminación de oración).
- 25 **[0207]** La unidad de cálculo de rasgo acústico 344b ingresa el rasgo de lenguaje de trama y las etiquetas de estilo de habla (la información de entonación acentuada y la información de terminación de oración) a cada unidad de la capa de entrada I del modelo de rasgo acústico 346b que se muestra en la Fig. 17, por ejemplo. La unidad de cálculo de rasgo acústico 344b calcula la DNN usando los parámetros del modelo de rasgo acústico aprendido 346b, y establece la salida de valor de la capa de salida O como el rasgo acústico.
- 30 **[0208]** La unidad de cálculo de rasgo acústico 344b emite el rasgo acústico generado para cada trama a la unidad de síntesis 345.
- [0209]** La unidad de síntesis 345 realiza la síntesis de habla usando el rasgo acústico para cada trama generada por la unidad de cálculo de rasgo acústico 344b de la unidad de cálculo de DNN 344, y genera habla sintetizada.
- 35 **[0210]** La unidad de síntesis 345 genera una forma de onda de fuente de sonido de cuerda vocal que expresa el cambio temporal del tono, la sonoridad y el nivel de ruido de la voz, en función de información (parámetros de fuente de sonido) tal como las características de tono y ruido incluidas en el rasgo acústico. La unidad de síntesis 345 genera un filtro de tracto vocal que expresa diferencias en fonemas y calidad de voz, en función de información (parámetros de espectro) tal como un coeficiente de espectro incluido en el rasgo acústico para cada trama. La unidad de síntesis 345 genera una forma de onda del habla correspondiente a la trama, utilizando la forma de onda de fuente de sonido de cuerda vocal como una entrada para el filtro de tracto vocal.
- 40 **[0211]** Dado que como procedimiento para realizar la síntesis de habla usando un rasgo acústico, se puede usar un procedimiento de vocoder conocido públicamente, se omite aquí una descripción detallada del mismo.
- 45 **[0212]** La unidad de síntesis 345 genera un habla sintetizada al conectar la forma de onda del habla de cada trama y emite datos de descripción de audio que indican el habla sintetizada generada.
- [0213]** Al configurar la unidad de generación de habla sintetizada 341 como se describió anteriormente, el dispositivo de generación de orientación de audio 30 utiliza el modelo de audio preaprendido por la unidad de preaprendizaje 31 puede generar, a partir de los datos de la competición, un habla sintetizada que expresa la entonación acentuada y termina una oración con una frase nominal que se puede utilizar como un deporte en vivo.
- 50 **[0214]** Por ejemplo, cuando la síntesis de habla se realiza sin aprender entonación acentuada, como se muestra en la Fig. 25, el audio que tiene un tono plano generalmente se sintetiza como un discurso que representa «KONO SETTO 6 KAIME NO BUREEKU». Además, cuando la síntesis de habla se realiza sin aprender la terminación de una oración con una frase nominal, como se muestra en la Fig. 25, generalmente se sintetiza un habla de tono descendente, como un habla que representa «BUREEKU» al final de una oración. En la Fig. 25, se muestra un tono de voz general L en la cadena de caracteres.
- 60 **[0215]** Por otro lado, dado que el dispositivo de generación de orientación de audio 30 utiliza un modelo de audio obtenido mediante el aprendizaje de la entonación acentuada y la terminación de una oración con una frase nominal, como se muestra en la Fig. 26, el tono del habla que representa el «6 KAIME NO BUREEKU», que es la porción acentuada, aumenta, y el tono del habla que representa el «BUREEKU» del final de la oración, que es la terminación de una oración con una frase nominal, se eleva.
- 65

5 **[0216]** La configuración del dispositivo de generación de orientación de audio 30 según la realización de la presente invención se ha descrito anteriormente. Sin embargo, el dispositivo de generación de orientación de audio 30 puede ser operado por un programa (programa de síntesis de habla) para hacer que un ordenador funcione como cada unidad descrita anteriormente. (Funcionamiento del dispositivo de generación de orientación de audio)

10 **[0217]** A continuación, con referencia a la Fig. 27 y la Fig. 28, se describirá el funcionamiento del dispositivo de generación de orientación de audio 30 según la realización de la presente invención. Aquí, el funcionamiento del dispositivo de generación de orientación de audio 30 se describirá por separado para el funcionamiento de la unidad de preaprendizaje 31 y el funcionamiento de la unidad de generación de habla sintetizada 341. (Preaprendizaje)

[0218] En primer lugar, con referencia a la Fig. 27, se describirá el funcionamiento de la unidad de preaprendizaje 31 del dispositivo de generación de orientación de audio 30.

15 **[0219]** (Etapa S51) La unidad de entrada de corpus 32 recibe un corpus del dispositivo de almacenamiento de corpus 40, con el texto y el audio (audio de entonación acentuada o audio normal) como un par. Aquí, la unidad de entrada de corpus 32 agrega información de discriminación de entonación acentuada que indica que el texto corresponde al audio de entonación acentuada, al texto correspondiente al audio de entonación acentuada. La unidad de entrada de corpus 32 agrega información de discriminación de entonación acentuada que indica que el texto no es
20 el audio de entonación acentuada, al texto correspondiente al audio normal.

[0220] (Etapa S52) La unidad de generación de rasgo de lenguaje de fonema 331a de la unidad de análisis de texto 331 analiza la entrada de texto en la etapa S51 y genera un rasgo de lenguaje (rasgo de lenguaje de fonema) para cada fonema que constituye el texto.
25

[0221] (Etapa S53) La unidad de generación de rasgo de lenguaje de fonema 331a agrega una etiqueta de estilo de habla al rasgo de lenguaje de fonema generado en la etapa S52. En un caso en el que la información de discriminación de entonación acentuada que indica que el texto es audio de entonación acentuada se agrega al texto, la unidad de generación de rasgo de lenguaje de fonema 331a agrega, como etiqueta de estilo de habla, una etiqueta
30 que indica la entonación acentuada (información de entonación acentuada; con entonación acentuada). En un caso en el que la información de discriminación de entonación acentuada que indica que el texto no es el audio de entonación acentuada se agrega al texto, la unidad de generación de rasgo de lenguaje de fonema 331a agrega, como etiqueta de estilo de habla, una etiqueta que indica normal (información de entonación acentuada; sin entonación acentuada).
35

[0222] La unidad de generación de rasgo de lenguaje de fonema 331a agrega, como la etiqueta de estilo de habla, una etiqueta que indica que termina una oración con una frase nominal (información de terminación de oración; con la terminación de una oración con una frase nominal), al rasgo de lenguaje de fonema, con respecto a los fonemas correspondientes a palabras de cual parte del habla de las palabras de terminación de oración son sustantivos. La
40 unidad de generación de rasgo de lenguaje de fonema 331a agrega, como la etiqueta de estilo de habla, una etiqueta que indica que no es la terminación de una oración con una frase nominal (información de terminación de oración; sin terminar una oración con una frase nominal), al rasgo de lenguaje de fonema, con respecto a los fonemas correspondientes a palabras que no sean las palabras de terminación de oración, o palabras que son palabras de terminación de oración pero no son sustantivos.
45

[0223] (Etapa S54) El separador de fonemas 332a de la unidad de análisis de audio 332 detecta la posición de ruptura del fonema de la entrada de habla en la etapa S51 y obtiene la duración del tiempo del fonema.

50 **[0224]** (Etapa S55) La unidad de aprendizaje de modelo de duración del tiempo 333a de la unidad de aprendizaje de DNN 333 recibe el rasgo de lenguaje de fonema generado en la etapa S52 y la etiqueta de estilo de habla agregada en la etapa S53, y aprende el modelo de duración del tiempo 346a de modo que la duración del tiempo del fonema obtenido en la etapa S54 se va a emitir.

55 **[0225]** (Etapa S56) La unidad de generación de rasgo de lenguaje de trama 331b de la unidad de análisis de texto 331 genera el rasgo de lenguaje (rasgo de lenguaje de trama) de los tramas que constituyen la sección de la duración del tiempo de fonema, en función de la duración del tiempo de fonema obtenida en la etapa S54.

60 **[0226]** (Etapa S57) La unidad de generación de rasgo de lenguaje de trama 331b agrega la misma etiqueta de estilo de habla que la etiqueta de estilo de habla agregada al rasgo de lenguaje de fonema en la etapa S53 al rasgo de lenguaje de trama generado en la etapa S56.

[0227] (Etapa S58) La unidad de generación de rasgo acústico 332b de la unidad de análisis de audio 332 genera un rasgo acústico de la trama en función de la posición de ruptura del fonema obtenida en la etapa S54 para la entrada de audio en la etapa S51.
65

[0228] (Etapa S59) La unidad de aprendizaje de modelo de rasgo acústico 333b de la unidad de aprendizaje de DNN 333 recibe el rasgo de lenguaje de trama generado en la etapa S56 y la etiqueta de estilo de habla agregada en la etapa S57, y aprende el modelo de rasgo acústico 346b de modo que el rasgo acústico generado en la etapa S58 se va a emitir. A continuación, la unidad de preaprendizaje 31 finaliza el procedimiento de la Fig. 27.

5 (Procesamiento de generación de descripción de audio)

[0229] A continuación, con referencia a la Fig. 28, se describirá la operación relacionada con el procesamiento de generación de descripción de audio del dispositivo de generación de orientación de audio 30.

10 **[0230]** (Etapa S60) La unidad de gestión de mensajes 11 recibe datos de competición desde el exterior del dispositivo de generación de orientación de audio 30.

15 **[0231]** (Etapa S61) La unidad de generación de explicación 12 genera un texto con el estilo de habla a partir de datos de competición no enviados entre los datos de competición introducidos en la etapa S60, en función de la plantilla explicativa almacenada de antemano en una unidad de almacenamiento de plantilla (no mostrada). En esta plantilla, las etiquetas predeterminadas (<entonación acentuada>, </entonación acentuada>) se agregan a las ubicaciones donde se debe pronunciar el habla con entonación acentuada (véase la Fig. 23). La unidad de generación de explicación 12 genera un texto con un estilo de habla, reemplazando la parte variable de la plantilla explicativa con
20 información especificada por los datos de la competición (véase la Fig. 23).

[0232] (Etapa S62) La unidad de análisis de estilo 342 de la unidad de generación de habla sintetizada 341 analiza el texto con el estilo de habla generado en la etapa S61, y genera el texto del que se ha eliminado la etiqueta y la información de posición de entonación acentuada que indica la porción acentuada en el texto designado por la
25 etiqueta.

[0233] (Etapa S63) La unidad de generación de rasgo de lenguaje de fonema 343a de la unidad de análisis de texto 343 analiza el texto generado en la etapa S62 y genera un rasgo de lenguaje (rasgo de lenguaje de fonema) para cada fonema que constituye el texto.

30 **[0234]** (Etapa S64) La unidad de generación de rasgo de lenguaje de fonema 343a agrega una etiqueta de estilo de habla al rasgo de lenguaje de fonema generado en la etapa S63. Aquí, la unidad de generación de rasgo de lenguaje de fonema 343a agrega, como una etiqueta de estilo de habla, una etiqueta que indica la entonación acentuada (información de entonación acentuada; con entonación acentuada), a los fonemas correspondientes a la
35 porción acentuada para la que se designa la entonación acentuada, en función de la información de posición de entonación acentuada, y añade una etiqueta que indica que no es la entonación acentuada (información de entonación acentuada; sin entonación acentuada) a los otros fonemas.

40 **[0235]** Con respecto a los fonemas correspondientes a palabras cuya parte del habla de las palabras de terminación de oración son sustantivos, la unidad de generación de rasgo de lenguaje de fonema 343a agrega, como etiqueta de estilo de habla, una etiqueta que indica que termina una oración con una frase nominal (información de terminación de oración; con la terminación de una oración con una frase nominal), al rasgo de lenguaje de fonema. Con respecto a los fonemas correspondientes a palabras que no sean las palabras de terminación de oración, o palabras que son palabras de terminación de oración pero no son sustantivos, la unidad de generación de rasgo de
45 lenguaje de fonema 343a agrega, como etiqueta de estilo de habla, una etiqueta que indica no terminar una oración con una frase nominal (información de terminación de oración; sin terminar una oración con una frase nominal), al rasgo de lenguaje de fonema.

50 **[0236]** (Etapa S65) La unidad de cálculo de duración del tiempo 344a de la unidad de cálculo de DNN 344 utiliza el modelo de duración del tiempo aprendido 346a para generar una duración del tiempo del fonema, a partir del rasgo de lenguaje de fonema generado en la etapa S63 y la etiqueta de estilo de habla agregada en la etapa S64.

[0237] (Etapa S66) La unidad de generación de rasgo de lenguaje de trama 343b de la unidad de análisis de texto 343 genera el rasgo de lenguaje (rasgo de lenguaje de trama) de las tramas que constituyen la sección de la
55 duración del tiempo de fonema, en función de la duración del tiempo de fonema generada en la etapa S65.

[0238] (Etapa S67) La unidad de generación de rasgo de lenguaje de trama 343b agrega la misma etiqueta de estilo de habla que la etiqueta de estilo de habla agregada al rasgo de lenguaje de fonema en la etapa S64 al rasgo de lenguaje de trama generado en la etapa S66.

60 **[0239]** (Etapa S68) La unidad de cálculo de rasgo acústico 344b de la unidad de cálculo de DNN 344 utiliza el modelo de rasgo acústico aprendido 346b para generar el rasgo acústico de la trama a partir del rasgo de lenguaje de trama generado en la etapa S66 y la etiqueta de estilo de habla agregada en la etapa S67.

65 **[0240]** (Etapa S69) La unidad de síntesis 345 realiza la síntesis de habla usando el rasgo acústico de la trama generada en la etapa S68, y genera datos de descripción de audio que indican la voz sintetizada generada.

Posteriormente, el dispositivo de generación de orientación de audio 30 finaliza el procedimiento de la Fig. 28.

5 **[0241]** Con la operación descrita anteriormente, el dispositivo de generación de orientación de audio 30 puede generar un habla sintetizada que expresa la entonación acentuada y termina una oración con una frase nominal que se puede utilizar como un deporte en vivo, a partir de los datos de la competición.

10 **[0242]** Tal como se describió anteriormente, el dispositivo de generación de orientación de audio 30 según la presente realización es un dispositivo de generación de orientación de audio 30 que recibe un rasgo de lenguaje para cada fonema e información de entonación acústica que indica la presencia o ausencia de la entonación acústica en el fonema mediante el uso de una red neuronal profunda, y sintetiza el audio del contenido de la competición a partir de los datos de la competición, mediante el uso de un modelo de duración del tiempo preaprendido para generar una duración del tiempo del fonema, y un modelo de rasgo acústico preaprendido que recibe un rasgo de lenguaje para cada trama del fonema, e información de entonación acústica que indica la presencia o ausencia de la entonación acústica en la trama, y emite el rasgo acústico de la trama, y el dispositivo de generación de orientación de audio 30 incluye una unidad de generación de explicación 12, una unidad de generación de rasgo de lenguaje de fonema 343a, una unidad de cálculo de duración del tiempo 344a, una unidad de generación de rasgo de lenguaje de trama 343b, una unidad de cálculo de rasgo acústico 344b y una unidad de síntesis 345.

20 **[0243]** En dicha configuración, el dispositivo de generación de orientación de audio 30 utiliza la unidad de generación de explicación 12 para generar un texto que indica la porción acentuada de los datos de competición, utilizando una plantilla de texto en la que la porción acentuada está predeterminada para los datos de competición.

25 **[0244]** En el dispositivo de generación de orientación de audio 30, la unidad de generación de rasgo de lenguaje de fonema 343a genera un rasgo de lenguaje para cada fonema e información de entonación acentuada en el fonema, mediante la realización de un análisis morfológico, estimación prosódica y similares en el texto generado por la unidad de generación de explicación 12.

30 **[0245]** En el dispositivo de generación de orientación de audio 30, la unidad de cálculo de duración del tiempo 344a calcula el rasgo de lenguaje de fonema y la información de entonación acentuada generada por la unidad de generación de rasgo de lenguaje de fonema 343a como la entrada del modelo de duración del tiempo para generar una duración del tiempo del fonema.

35 **[0246]** En el dispositivo de generación de orientación de audio 30, la unidad de generación de rasgo de lenguaje de trama 343b genera un rasgo de lenguaje para cada trama del fonema, a partir de la duración del tiempo del fonema generado por la unidad de cálculo de duración del tiempo 344a y el rasgo de lenguaje del fonema generado por la unidad de generación de rasgo de lenguaje de fonema 343a, y asocia el rasgo de lenguaje generado con la información de entonación acentuada del fonema correspondiente a la trama. Por lo tanto, el rasgo de lenguaje y la información de entonación acentuada para cada trama que constituye la sección de duración del tiempo del fonema están asociadas entre sí.

40 **[0247]** En el dispositivo de generación de orientación de audio 30, la unidad de cálculo de rasgo acústico 344b calcula el rasgo de lenguaje de trama generado por la unidad de generación de rasgo de lenguaje de trama 343b y la información de entonación acentuada correspondiente a la trama como la entrada del modelo de rasgo acústico para generar el rasgo acústico para cada trama.

45 **[0248]** El dispositivo de generación de orientación de audio 30 utiliza la unidad de síntesis 345 para realizar la síntesis de habla mediante el uso del rasgo acústico para cada trama generada por la unidad de cálculo de rasgo acústico 344b.

50 **[0249]** Por lo tanto, el dispositivo de generación de orientación de audio 30 puede generar un habla sintetizada en la que las porciones acentuadas en el texto se acentúan usando un modelo de rasgo acústico en el que los rasgos del audio de entonación acentuada se aprenden de antemano.

55 **[0250]** El dispositivo de generación de orientación de audio 30 puede ser operado por un programa de generación de descripción de audio para hacer que un ordenador funcione como cada unidad del dispositivo de generación de orientación de audio 30.

60 **[0251]** Según la presente realización, es posible generar un texto que especifique una parte a destacar, a partir de los datos de la competición, y generar un habla sintetizada en la que la porción se acentúe.

[0252] Por lo tanto, en la presente realización, es posible generar audio que pueda expresar una intención con una inflexión clara, a partir de los datos de la competición.

65

(Ejemplo de modificación)

[0253] Aunque la configuración y el funcionamiento del dispositivo de generación de orientación de audio 30 según las realizaciones de la presente invención se ilustran en lo anterior, la presente invención no se limita a este ejemplo.

5 **[0254]** El dispositivo de generación de orientación de audio 30 realiza dos operaciones de preaprendizaje para aprender un modelo de audio (el modelo de duración del tiempo 346a y el modelo de rasgo acústico 346b) y procesamiento de síntesis para sintetizar habla a partir de datos de competición usando el modelo de audio, por un dispositivo. Sin embargo, estas operaciones pueden ser realizadas por dispositivos separados.

10 **[0255]** Específicamente, la unidad de preaprendizaje 31 que se muestra en la Fig. 13 se puede separar como un dispositivo de aprendizaje de modelo de audio independiente, como un dispositivo que preaprende un modelo de audio. En este caso, el dispositivo de aprendizaje de modelo de audio (la unidad de preaprendizaje 31) puede operarse con un programa (programa de aprendizaje de modelo de audio) que hace que el ordenador funcione como cada unidad del dispositivo de aprendizaje de modelo de audio descrito en la unidad de preaprendizaje 31.

15 **[0256]** En el aparato de generación de modelo de audio, la unidad de aprendizaje de modelo 33 que aprende el modelo de duración del tiempo y el modelo de rasgo acústico de una red neuronal profunda utilizada en la unidad de generación de habla sintetizada 341, usando un texto, un habla pronunciada con entonación acentuada en el texto y un habla pronunciada sin entonación acentuada, se configura para incluir una unidad de generación de rasgo de lenguaje de fonema 331a, un separador de fonemas 332a, una unidad de aprendizaje de modelo de duración del tiempo 333a, una unidad de generación de rasgo de lenguaje de trama 331b, una unidad de generación de rasgo acústico 332b y una unidad de aprendizaje de modelo de rasgo acústico 333b.

20

[0257] En dicha configuración, la unidad de generación de rasgo de lenguaje de fonema 331a realiza análisis tales como análisis morfológico y estimación de prosodia en el texto, y genera rasgo de lenguaje para cada fonema e información de entonación acentuada que indica la presencia o ausencia de la entonación acentuada para el fonema, dependiendo de si el texto corresponde al audio acentuado.

25

[0258] El separador de fonemas 332a detecta la ruptura para cada fonema del audio correspondiente al texto, y especifica la duración del tiempo del fonema.

30

[0259] La unidad de aprendizaje del modelo de duración del tiempo 333a aprende un modelo de duración del tiempo para recibir el rasgo de lenguaje de fonema y la información de entonación acentuada generada por la unidad de generación de rasgo de lenguaje de fonema 331a y emitir la duración del tiempo de fonema especificada por el separador de fonema 332a.

35

[0260] La unidad de generación de rasgo de lenguaje de trama 331b genera un rasgo de lenguaje para cada trama de fonema, a partir de la duración del tiempo de fonema especificada por el separador de fonema 332a y el rasgo de lenguaje generado por la unidad de generación de rasgo de lenguaje de fonema 331a, y asocia el rasgo de lenguaje generado con la información de entonación acentuada de fonema correspondiente a la trama.

40

[0261] La unidad de generación de rasgo acústico 332b genera un rasgo acústico para cada trama que constituye una sección de la duración del tiempo de fonema del audio correspondiente al texto, en función de la ruptura de fonema detectada por el separador de fonemas 332a.

45

[0262] La unidad de aprendizaje de modelo de rasgo acústico 333b aprende un modelo de rasgo acústico para recibir el rasgo de lenguaje de trama generado por la unidad de generación de rasgo de lenguaje de trama 331b y la información de entonación acentuada y emitir el rasgo acústico generado por la unidad de generación de rasgo acústico 332b.

50

[0263] Por lo tanto, el dispositivo de aprendizaje de modelo de audio puede generar un modelo que aprende una duración del tiempo y un rasgo acústico de un fonema que difiere dependiendo de la información de entonación acentuada, incluso si el texto es el mismo.

55 **[0264]** Por otro lado, la unidad de preaprendizaje 31 se puede omitir del dispositivo de generación de orientación de audio 30 ilustrado en la Fig. 13 como un dispositivo que realiza la síntesis de habla a partir de los datos de la competición. En este caso, el dispositivo de generación de orientación de audio 30 puede hacer que el ordenador opere con un programa (programa de generación de descripción de audio) para ejecutar el procesamiento de cada unidad de la unidad de gestión de mensajes 11, la unidad de generación de explicación 12 y la unidad de síntesis de habla 34.

60

[0265] De esta manera, dado que la operación de preaprendizaje para aprender el modelo de audio y la operación de procesamiento de síntesis desde la recepción de datos de competición hasta la síntesis de habla usando el modelo de audio se operan en diferentes dispositivos, el modelo de audio obtenido mediante el aprendizaje con un dispositivo de aprendizaje de modelo de audio (unidad de preaprendizaje 31) puede ser utilizado por una pluralidad

65

de dispositivos de generación de orientación de audio 30.

[0266] En la presente, se ha descrito que el dispositivo de generación de orientación de audio 30 genera un habla sintetizada que realiza la entonación acentuada y termina una oración con una frase nominal como un estilo del habla, pero la presente invención no se limita a esta.

[0267] Por ejemplo, el dispositivo de generación de orientación de audio 30 puede generar un habla sintetizada que solo realiza la entonación acentuada como un estilo del habla.

[0268] En este caso, como el modelo de duración del tiempo 346a y el modelo de rasgo acústico 346b, se puede usar un modelo en el que la información de terminación de oración (etiqueta que indica la presencia o ausencia de terminación de una oración con una frase nominal) se omite de la capa de entrada I del modelo de duración del tiempo 346a que se muestra en la Fig. 16 y la información de terminación de oración se omite de la capa de entrada I del modelo de rasgo acústico 346b que se muestra en la Fig. 17. La información de terminación de oración se puede omitir de la salida de etiqueta de estilo de habla de la unidad de análisis de texto 331 a la unidad de aprendizaje de DNN 333. La información de terminación de oración se puede omitir de la salida de etiqueta de estilo de habla de la unidad de análisis de texto 343 a la unidad de cálculo de DNN 344.

<Tercera realización

[0269] A continuación, se describirá una tercera realización de la presente invención. La siguiente descripción se centra principalmente en las diferencias de las realizaciones descritas anteriormente. Los mismos elementos y procedimientos constituyentes que en las realizaciones anteriores se indican con los mismos números de referencia, y se utiliza la descripción.

[0270] Cuando se proporciona un programa de difusión en vivo tal como una competición deportiva, el sistema de difusión descrito anteriormente 1 puede recibir tanto el flujo de distribución como los datos de la competición por medios o vías separadas. El flujo de distribución incluye datos de vídeo que representan regularmente la situación de la competición y datos de audio en vivo que representan el sonido de fondo. Los datos de la competición se pueden distribuir esporádicamente según la ocurrencia de una competición o un evento fuera de la competición. Se produce un retraso en la transmisión en la vía de transmisión donde estos datos se distribuyen desde el lugar de competición hasta la ubicación de instalación del equipo del sistema de difusión 1. La cantidad de retardo y las características de fluctuación del retardo de transmisión dependen de la vía de transmisión.

[0271] Incluso si tanto el flujo de distribución como los datos de competición se distribuyen de forma síncrona a través de una única línea de transmisión, un tiempo de procesamiento desde que se recibe el flujo de distribución hasta que se puede enviar al dispositivo receptor 20 es generalmente diferente de un tiempo de procesamiento requerido para generar datos de descripción de audio basados en los datos de competición y multiplexarlos con el flujo de distribución que incluye datos de audio en vivo que indican el sonido de fondo después de recibir los datos de competición. Dado que el tiempo requerido para el procesamiento de síntesis de habla depende de la longitud de la expresión, el tiempo de procesamiento no es constante y puede variar.

[0272] Por estas razones, en el sistema de difusión 1, la ocurrencia de un evento que aparece en el vídeo o sonido de fondo proporcionado al dispositivo receptor 20 y el audio de orientación indicado por los datos de descripción de audio generados en función de los datos de competición puede no sincronizarse. Por lo tanto, el audio de orientación puede retrasarse con respecto al evento. En algunos casos, el audio de orientación puede proporcionarse antes del vídeo que representa el evento pronunciado o el sonido de fondo.

[0273] Dado que los tiempos de retraso relacionados con las distribuciones de los datos de la competición pueden ser diferentes individualmente, cuando el funcionamiento del medio de distribución y la vía de transmisión son inestables, pueden recibirse en un orden diferente a la ocurrencia del evento. Si los datos de descripción de audio se generan secuencialmente en función de los datos de la competición y se distribuyen tal como están, los datos de descripción de audio se proporcionan en un orden diferente del orden en el que se han producido los eventos originalmente. Esto también causa una diferencia con el orden de los eventos indicados por el flujo de distribución formado por datos de vídeo y datos de audio originales.

[0274] Por lo tanto, como se ilustra en la Fig. 29, el sistema de difusión 5 según la presente realización incluye un dispositivo de generación de orientación de audio 50, una unidad de procesamiento de vídeo 55 y un dispositivo transmisor 56.

[0275] El dispositivo de generación de orientación de audio 50 incluye una unidad de gestión de mensajes 11, una unidad de generación de explicación 12, una unidad de síntesis de habla 13 y una primera unidad de recepción 17.

[0276] El dispositivo transmisor 56 incluye una primera unidad de búfer 561, una segunda unidad de búfer 562

y una unidad de multiplexación 563. La primera unidad de búfer 561 almacena los datos de descripción de audio y la información de tiempo recibida secuencialmente del dispositivo de generación de orientación de audio 10 en cada unidad predeterminada en asociación entre sí. La unidad predeterminada es, por ejemplo, una trama que tiene una duración del tiempo predeterminada (por ejemplo, 20 ms a ls), y puede ser un período igual o menor que el intervalo de ocurrencia del evento. La información de tiempo es información sobre el momento en que se ha producido la situación de la competición indicada por los datos de competición utilizados por la unidad de generación de explicación 12 para generar el texto explicativo ha medida que la información no transmitida se ha producido. Los datos de la competición se adquieren en asociación con la información de tiempo. El texto explicativo es utilizado por la unidad de síntesis de habla 13 para generar datos de descripción de audio después de realizar el procesamiento de síntesis de habla.

[0277] La segunda unidad de búfer 562 almacena la entrada de datos de vídeo de la unidad de procesamiento de vídeo 55 y la información de tiempo en asociación entre sí. Los datos de vídeo son datos que indican un vídeo de la situación de la competición. La información de tiempo es información que indica el momento en que ocurre la situación. Los datos de vídeo pueden estar asociados con datos de audio originales que indican el audio de la situación de la competición y almacenados en la segunda unidad de búfer 562. En la siguiente descripción, los datos de vídeo y los datos de audio originales se denominan colectivamente flujo de distribución.

[0278] La unidad de multiplexación 563 lee los datos de descripción de audio en el tiempo t-D (en lo sucesivo denominado tiempo designado) que es un tiempo de desplazamiento predeterminado D antes del tiempo actual y el flujo de distribución en el tiempo t-D de la primera unidad de búfer 561 y la segunda unidad de búfer 562, respectivamente. En la unidad de multiplexación 563, un tiempo de desplazamiento D se establece de antemano. El tiempo de desplazamiento D puede ser un tiempo mayor que el tiempo requerido desde la recepción de datos de competición hasta el inicio de la transmisión de datos de descripción de audio en el dispositivo de generación de orientación de audio 50. La unidad de multiplexación 563 multiplexa los datos de descripción de audio leído y el flujo de distribución para generar datos multiplexados para su distribución al dispositivo receptor 20.

[0279] Más específicamente, cuando los datos de audio originales no se incluyen en el flujo de distribución y se incluyen los datos de vídeo, la unidad de multiplexación 563 multiplexa los datos de descripción de audio y los datos de vídeo para generar datos multiplexados. Cuando los datos de audio originales se incluyen en el flujo de distribución, la unidad de multiplexación 563 puede mezclar el audio de orientación indicado por los datos de descripción de audio y el audio indicado por los datos de audio originales. La unidad de multiplexación 563 genera datos multiplexados al multiplexar los datos de audio mezclados que indican el audio mezclado obtenido por la mezcla y los datos de vídeo incluidos en el flujo de distribución. La unidad de multiplexación 563 transmite los datos multiplexados generados a la vía de transmisión de difusión o la vía de transmisión de comunicación.

[0280] El dispositivo receptor 20 recibe los datos multiplexados a través de la vía de transmisión de difusión o la vía de transmisión de comunicación, y separa los datos multiplexados recibidos en datos de vídeo y datos de audio mixtos o datos de descripción de audio. El dispositivo receptor 20 muestra vídeo en función de los datos de vídeo separados, y reproduce audio en función de los datos de audio mixtos o datos de descripción de audio.

[0281] El dispositivo de generación de orientación de audio 10 incluye además una primera unidad de recepción 17. La primera unidad de recepción 17 recibe, por ejemplo, datos de competición que indican la última situación de la competición e información de tiempo que indica el momento en que la situación ha ocurrido, desde el centro de datos, y lo emite a la unidad de gestión de mensajes 11. La primera unidad de recepción 17 incluye, por ejemplo, una interfaz de comunicación.

[0282] La unidad de gestión de mensajes 11 almacena la entrada de datos de competición de la primera unidad de recepción 17 y la información de tiempo en asociación entre sí. La información de tiempo puede describirse en los datos de la competición. En ese caso, la primera unidad de recepción 17 no recibe información de tiempo por separado de los datos de competición, y la unidad de gestión de mensajes 11 puede extraer la información de tiempo descrita en los datos de competición.

[0283] La unidad de gestión de mensajes 11 puede recibir además una señal de sincronización que indica información de tiempo que indica el tiempo desde el centro de datos de la fuente de transmisión a través de la primera unidad de recepción 17, y sincronizarse con el centro de datos en función de la información de tiempo recibida.

[0284] Desde la unidad de gestión de mensajes 11, la unidad de generación de explicación 12 se refiere a los datos de competición en el tiempo t-D' que es el segundo tiempo de desplazamiento predeterminado D' antes del tiempo actual t, como información no transmitida, entre los datos de competición acumulados en la unidad de gestión de mensajes 11. En la unidad de generación de explicación 12, un segundo tiempo de desplazamiento D' se establece de antemano. El segundo tiempo de desplazamiento D' es un tiempo mayor que la cantidad de fluctuación estándar del tiempo de transmisión desde el centro de datos al dispositivo de generación de orientación de audio 10, y puede ser un valor positivo más corto que el primer tiempo de desplazamiento D ($0 < D' < D$). La unidad de generación de explicación 12 utiliza los datos de competición referenciados para generar un texto explicativo. En otras palabras, es

deseable que el primer tiempo de desplazamiento D sea igual o superior al tiempo obtenido mediante la adición del valor máximo del tiempo requerido para que la unidad de generación de explicación 12 genere el texto explicativo a partir de la información no transmitida y el valor máximo del tiempo requerido para generar los datos de descripción de audio del texto explicativo al segundo tiempo de desplazamiento D'. Si la diferencia D-D' entre el tiempo de desplazamiento D y el segundo tiempo de desplazamiento D' es menor que el tiempo de procesamiento Δ , que es la suma del tiempo requerido para generar el texto explicativo y el tiempo requerido para generar los datos de descripción de audio, la unidad de multiplexación 563 puede retrasar los datos de distribución por la diferencia de tiempo $\Delta - D + D'$ y multiplexarlos con los datos de descripción de audio. En general, el tiempo necesario para generar el texto explicativo y el tiempo necesario para generar los datos de audiodescripción tienden a aumentar a medida que el texto explicativo y los datos de audiodescripción se hacen más largos. Cuando el valor máximo del tiempo de retraso permitido para la unidad de multiplexación 563 se establece como el tiempo de desplazamiento D, en un caso donde el tiempo de desplazamiento D es menor que el tiempo de procesamiento Δ que es la suma de estos tiempos, el audio de orientación en función de los datos de descripción de audio se retrasa. En general, el audio de orientación más corto tiende a tener una respuesta rápida a su contenido, y el audio de orientación más largo tiende a tener poca respuesta rápida a su contenido. Por lo tanto, tal retraso puede permitirse incluso si ocurre temporalmente.

[0285] La unidad de procesamiento de vídeo 55 incluye una segunda unidad de recepción 551. La segunda unidad de recepción 551 recibe, por ejemplo, un flujo de suministro que indica una situación de la competición de un centro de datos, e información de tiempo que indica un momento en que la situación ha ocurrido, en una unidad predeterminada. La segunda unidad de recepción 551 almacena secuencialmente el flujo de distribución recibido y la información de tiempo en asociación entre sí en la segunda unidad de búfer 562 del dispositivo transmisor 56. La unidad predeterminada es, por ejemplo, una trama de vídeo. Como información de tiempo, por ejemplo, se puede usar una marca de tiempo que indique la hora en que se genera el vídeo. La unidad predeterminada puede ser un período igual o menor que el intervalo de ocurrencia del evento. La segunda unidad de recepción 551 incluye, por ejemplo, una interfaz de comunicación.

[0286] La primera unidad de búfer 561 y la segunda unidad de búfer 562 pueden configurarse cada una como un búfer First-in First-out (FIFO). Según la configuración, en la primera unidad de búfer 561, se forma una primera cola FIFO como una serie de tiempo de un conjunto que incluye datos de descripción de audio e información de tiempo. En la segunda unidad de búfer 562, se forma una segunda cola FIFO como una serie de tiempo de un conjunto que incluye datos de distribución e información de tiempo.

[0287] A continuación, se describirá el procesamiento de distribución de descripción de audio según la presente realización.

[0288] La Fig. 30 es un diagrama de flujo que ilustra un ejemplo de procesamiento de distribución de descripción de audio según la presente realización.

[0289] (Etapa S71) La segunda unidad de recepción 551 recibe el flujo de distribución que indica la situación de la competición, y la información de tiempo, en unidades predeterminadas. La segunda unidad de recepción 551 almacena secuencialmente el flujo de distribución recibido y la información de tiempo en asociación entre sí en la segunda unidad de búfer 562 del dispositivo transmisor 56.

[0290] (Etapa S72) La primera unidad de recepción 17 recibe datos de competición que indican la última situación de la competición e información de tiempo que indica la hora en que ha ocurrido la situación. La unidad de gestión de mensajes 11 almacena los datos de competición y la información de tiempo, introducida desde la primera unidad de recepción 17, en asociación entre sí.

[0291] (Etapa S73) La unidad de generación de explicación 12 se refiere a los datos de competición en el tiempo t-D que es el segundo tiempo de desplazamiento D' antes del tiempo actual t, como información no transmitida, entre los datos de competición acumulados en la unidad de gestión de mensajes 11, de la unidad de gestión de mensajes 11. La unidad de generación de explicación 12 utiliza los datos de competición referenciados para generar un texto explicativo.

[0292] (Etapa S74) La unidad de síntesis de habla 13 realiza el procesamiento de síntesis de habla en el texto explicativo generado por la unidad de generación de explicación 12 para generar datos de descripción de audio, y almacena los datos de descripción de audio generados en asociación con información de tiempo que indica el tiempo relacionado con los datos de competición utilizados para generar el texto explicativo, en la primera unidad de búfer 561.

[0293] (Etapa S75) La unidad de multiplexación 563 lee los datos de descripción de audio en el tiempo designado t-D que es el primer tiempo de desplazamiento D antes del tiempo actual t de la primera unidad de búfer 561, y lee el flujo de distribución en el tiempo designado t-D de la segunda unidad de búfer 562. La unidad de multiplexación 563 genera datos multiplexados al multiplexar los datos de descripción de audio leídos y el flujo de distribución, y transmite los datos multiplexados generados a la vía de transmisión de difusión. Posteriormente, se

termina el procedimiento de la Fig. 30.

(Ejemplo de modificación)

- 5 **[0294]** El flujo de distribución no siempre se transmite en asociación con la información de tiempo. Por lo tanto, la segunda unidad de búfer 562 puede calcular el tiempo t' que indica la situación de la competición indicada por el flujo de distribución restando el tiempo de retraso δ desde el momento t cuando se adquiere el flujo de distribución. Por ejemplo, la segunda unidad de búfer 562 recibe la señal de sincronización de la fuente de transmisión (por ejemplo, el centro de datos) del flujo de distribución a través de la segunda unidad de recepción 551, y puede medir el tiempo
- 10 de retraso δ restando el tiempo t_1 indicado por la señal de sincronización recibida del tiempo actual t_0 . La segunda unidad de búfer 562 almacena el flujo de distribución adquirida en asociación con información de tiempo que indica el tiempo calculado t' . Por lo tanto, cuando la fluctuación de transmisión de la fuente de transmisión al dispositivo de generación de orientación de audio 50 es tan pequeña que se puede ignorar, el evento representado por el flujo de distribución se puede sincronizar con el audio de orientación indicado por los datos de descripción de audio generados
- 15 a partir de los datos de competición.
- [0295]** El dispositivo de generación de orientación de audio 50 puede incluir una unidad de síntesis de habla 34 (Fig. 13) en lugar de la unidad de síntesis de habla 13.
- 20 **[0296]** Tal como se describió anteriormente, según el sistema de difusión 5 según la presente realización, es posible sincronizar y distribuir los datos de distribución que indican una situación de la competición y los datos de descripción de audio relacionados con la situación. Dado que las fluctuaciones en el retraso de transmisión de los datos de competición se absorben, es posible realizar de manera más estable la sincronización entre los datos de descripción de audio generados en función de los datos de competición y los datos de distribución.
- 25 **[0297]** Tal como se describió anteriormente, las realizaciones de la invención se han descrito en detalle con referencia a los dibujos, pero la configuración específica no se limita a la descripción anterior, y se pueden realizar diversos cambios de diseño y similares sin apartarse del alcance de la presente invención.
- 30 **[0298]** Por ejemplo, el dispositivo de generación de orientación de audio 10, 30 puede estar integrado con una o ambas de la unidad de procesamiento de vídeo 15 y el dispositivo transmisor 16 y configurado como un único dispositivo de generación de orientación de audio
- 35 10, 30.
- [0299]** El dispositivo de generación de orientación de audio 50 puede estar integrado con una o ambas de la unidad de procesamiento de vídeo 55 y el dispositivo transmisor 56 y configurado como un único dispositivo de generación de orientación de audio 50.
- 40 **[0300]** El sistema de difusión 1, 5 puede incluir además un dispositivo servidor (no mostrado) y puede configurarse como un sistema de cooperación de comunicación de difusión. En ese caso, el dispositivo de generación de orientación de audio 10, 30 y 50 emite los datos de descripción de audio generados al dispositivo servidor. El dispositivo servidor puede transmitir datos de descripción de audio al dispositivo receptor 20 conectado a través de la vía de transmisión de comunicación. Por otro lado, el dispositivo receptor 20 proporciona la función del servicio de
- 45 cooperación de difusión de comunicación mediante la ejecución del procedimiento indicado por el comando descrito en el programa de aplicación. Como una función del dispositivo receptor 20, el dispositivo receptor 20 reproduce audio basándose en los datos de descripción de audio recibidos del dispositivo servidor.
- 50 **[0301]** El dispositivo de generación de orientación de audio 10, 30 puede no ser necesariamente parte del sistema de difusión 1. Por ejemplo, el dispositivo de generación de orientación de audio 10, 30 puede emitir los datos de descripción de audio a un dispositivo servidor conectado a la vía de transmisión de comunicación. El dispositivo servidor puede transmitir la entrada de datos de descripción de audio desde el dispositivo de generación de orientación de audio 10, 30 y los datos de vídeo de la competición en asociación entre sí a varios dispositivos terminales conectados a la vía de transmisión de comunicación. El dispositivo servidor puede transmitir los datos de vídeo y los
- 55 datos de descripción de audio como una respuesta a la recepción de la señal de solicitud de distribución desde el dispositivo terminal, o puede transmitir todos los datos a los dispositivos terminales registrados previamente.
- 60 **[0302]** El dispositivo transmisor 56 del sistema de difusión 5 puede transmitir los datos multiplexados generados a la transmisión de comunicación en lugar de con la vía de transmisión de difusión o junto con la vía de transmisión de difusión.
- [0303]** El dispositivo de generación de orientación de audio descrito anteriormente 10, 30 y 50 puede configurarse como hardware dedicado o puede configurarse como hardware que incluye un sistema informático en este. Los procedimientos de funcionamiento de una parte del dispositivo de generación de orientación de audio 10, 30, 50 se almacenan en un medio de grabación legible por ordenador en forma de un programa, y los procedimientos
- 65

descritos anteriormente se pueden realizar mediante la lectura y ejecución del programa por parte del sistema informático. El sistema informático aquí incluye un procesador tal como una unidad de procesamiento central (CPU), varias memorias, un sistema operativo (OS) y hardware tal como dispositivos periféricos. Cada unidad funcional descrita anteriormente puede configurarse mediante uno o dos o más sistemas informáticos, y un conjunto de una 5 unidad funcional o dos o más unidades funcionales cada una puede configurarse mediante un sistema informático.

[0304] Además, «medio de grabación legible por ordenador» se refiere a un medio portátil tal como un disco flexible, un disco magnetoóptico, una memoria de solo lectura (Rom) y un disco compacto (CD)-ROM, y un dispositivo de almacenamiento tal como un disco duro incorporado en un sistema informático. El «medio de grabación legible por 10 ordenador» puede incluir un medio que sostiene dinámicamente un programa durante un período de tiempo corto, tal como una línea de comunicación para transmitir un programa a través de una red tal como Internet o una línea de comunicación tal como una línea telefónica, y un medio que mantiene un programa durante un período de tiempo determinado, tal como una memoria volátil dentro de un sistema informático que sirve como un servidor o un cliente en ese caso. Además, el programa puede estar destinado a realizar una parte de las funciones descritas anteriormente, 15 y puede estar destinado a realizar las funciones descritas anteriormente combinándolas con el programa ya registrado en el sistema informático.

Lista de signos de referencia

20 **[0305]**

- 1, 5: sistema de difusión,
- 10, 30, 50: dispositivo de generación de orientación de audio,
- 11: unidad de gestión de mensajes,
- 25 12: unidad de generación de explicación,
- 13, 34: unidad de síntesis de habla,
- 15, 55: unidad de procesamiento de vídeo,
- 16, 56: dispositivo transmisor,
- 17: primera unidad de recepción,
- 30 20: dispositivo receptor,
- 31: unidad de preaprendizaje,
- 32: unidad de entrada de corpus,
- 33: unidad de aprendizaje de modelo,
- 34: unidad de síntesis de habla,
- 35 331: unidad de análisis de texto,
- 331a: unidad de generación de rasgo de lenguaje de fonema,
- 331b: unidad de generación de rasgo de lenguaje de trama,
- 332: unidad de análisis de audio,
- 332a: separador de fonemas,
- 40 332b: unidad de generación de rasgo acústico,
- 333: unidad de aprendizaje de DNN,
- 333a: unidad de aprendizaje del modelo de duración del tiempo,
- 333b: unidad de aprendizaje del modelo de rasgo acústico,
- 341: unidad de generación de habla sintetizada,
- 45 342: unidad de análisis de estilo,
- 343: unidad de análisis de texto,
- 343a: unidad de generación de rasgo de lenguaje de fonema,
- 343b: unidad de generación de rasgo de lenguaje de trama,
- 344: unidad de cálculo de DNN,
- 50 344a: unidad de cálculo de duración del tiempo,
- 344b: unidad de cálculo de rasgo acústico,
- 345: unidad de síntesis,
- 346: unidad de almacenamiento de modelo,
- 346a: modelo de duración del tiempo,
- 55 346b: modelo de rasgo acústico,
- 551: segunda unidad de recepción,
- 561: primera unidad de búfer,
- 562: segunda unidad de búfer,
- 563: unidad de multiplexación

REIVINDICACIONES

1. Un dispositivo de generación de orientación de audio (10, 30, 50) que comprende:

5 una unidad de gestión de mensajes (11) que recibe y acumula un mensaje, en la que el mensaje se distribuye para cada actualización, siendo el mensaje datos que representan una situación más reciente de una competición; una unidad de generación de explicación (12) que genera un texto explicativo para transmitir información no transmitida detectada a partir del mensaje, en función de la información transmitida; y
 10 una unidad de síntesis de habla (13, 34) que emite un habla convertida del texto explicativo, en la que la unidad de generación de explicación (12) almacena la información no transmitida para el texto explicativo como la información transmitida después de generar el texto explicativo, se mantiene en espera hasta que finalice la emisión del habla, e inicia un procedimiento para generar un nuevo texto explicativo en función de información no transmitida actualizada, **caracterizada porque**
 15 la unidad de gestión de mensajes (11) detecta una porción modificada del mensaje debido a la actualización del mensaje, genera nueva información según un tipo de la parte variada, y acumula aún más la nueva información generada.

20 2. El dispositivo de generación de orientación de audio (10, 30, 50) según la reivindicación 1, en el que la unidad de gestión de mensajes (11)

selecciona una plantilla de generación de información correspondiente al tipo de la porción modificada de plantillas de generación de información prealmacenadas, cada una de las plantillas de generación de información prealmacenadas indica una estructura de la nueva información para cada tipo de la porción modificada, y
 25 genera la nueva información mediante la integración de elementos de la parte variada en la plantilla de generación de información seleccionada.

3. El dispositivo de generación de orientación de audio (10, 30, 50) según la reivindicación 1 o 2, en el que la unidad de generación de explicación (12) considera que los tipos de información predeterminados, entre las piezas de información retenidas como la información transmitida, no se han transmitido, en un caso en el que ha transcurrido un tiempo predeterminado desde un tiempo en que se completa la transmisión.

4. El dispositivo de generación de orientación de audio (10, 30, 50) según cualquiera de las
 35 reivindicaciones 1 a 3, en el que la unidad de generación de explicación (12)

selecciona una plantilla explicativa correspondiente al tipo de la información no transmitida a partir de plantillas explicativas prealmacenadas, cada una de las plantillas explicativas prealmacenadas indica un patrón de oración del texto explicativo para cada tipo de información no transmitida, y
 40 genera el texto explicativo mediante la integración de elementos de la información no transmitida en la plantilla explicativa seleccionada.

5. El dispositivo de generación de orientación de audio (30) según cualquiera de las reivindicaciones 1 a
 45 4,

en el que la unidad de generación de explicación (12) utiliza una plantilla de texto en la que una porción acentuada está predeterminada para los datos de competición que indican una situación de la competición indicada por el mensaje para generar el texto explicativo que indica la porción acentuada de los datos de competición,
 50 en el que la unidad de síntesis de habla (34) comprende

una unidad de generación de rasgo de lenguaje de fonema (343a) que analiza el texto explicativo para generar un rasgo de lenguaje para cada fonema e información de entonación acentuada para el fonema;
 una unidad de cálculo de duración del tiempo (344a) que genera una duración del tiempo del fonema a partir
 55 del rasgo de lenguaje y la información de entonación acentuada generada por la unidad de generación de rasgo de lenguaje de fonema, utilizando un modelo de duración del tiempo;
 una unidad de generación de rasgo de lenguaje de trama (343b) que genera un rasgo de lenguaje para cada trama del fonema a partir de la duración del tiempo del fonema generado por la unidad de cálculo de duración del tiempo y el rasgo de lenguaje del fonema, y asocia el rasgo de lenguaje generado con la información de entonación acentuada del fonema correspondiente a la trama;
 60 una unidad de cálculo de rasgo acústico (344b) que genera un rasgo acústico para cada trama, a partir del rasgo de lenguaje de la trama generada por la unidad de generación de rasgo de lenguaje de trama y la información de entonación acentuada correspondiente a la trama, utilizando un modelo de rasgo acústico; y
 una unidad de síntesis (345) que sintetiza el habla usando el rasgo acústico para cada trama generada por la
 65 unidad de cálculo de rasgo acústico,

en la que el modelo de duración del tiempo se aprende preliminarmente, usando el rasgo de lenguaje para cada fonema, e información de entonación acústica que indica la presencia o ausencia de entonación acústica para el fonema como una entrada, y la duración del tiempo del fonema como una salida, y en la que el modelo de rasgo acústico se aprende preliminarmente, usando el rasgo de lenguaje para cada trama del fonema e información de entonación acústica que indica la presencia o ausencia de entonación acústica de la trama como una entrada, y la duración del tiempo del fonema como una salida, y un rasgo acústico de la trama como una salida.

6. El dispositivo de generación de orientación de audio (30) según la reivindicación 5,

en el que el modelo de duración del tiempo y el modelo de rasgo acústico se aprenden además usando información de terminación de oración que indica si el fonema es un fonema de una palabra de terminación de oración que tiene una parte del habla como un sustantivo,

en el que la unidad de generación de rasgo de lenguaje de fonema (343a) genera la información de terminación de oración para cada fonema, en función de la parte del habla de la palabra de terminación de oración del texto, en el que la unidad de cálculo de duración del tiempo (344a) genera una duración del tiempo del fonema, mediante el uso del modelo de duración del tiempo, a partir del rasgo de lenguaje del fonema, la información de entonación acentuada y la información de terminación de oración,

en el que la unidad de generación de rasgo de lenguaje de trama (343b) asocia el rasgo de lenguaje para cada trama con la información de terminación de oración, y

en el que la unidad de cálculo de rasgo acústico (344b) genera un rasgo acústico para cada trama, mediante el uso del modelo de rasgo acústico, a partir del rasgo de lenguaje de la trama, y la información de entonación acentuada y la información de terminación de oración correspondiente a la trama.

7. Un sistema de difusión (5) que comprende:

el dispositivo de generación de orientación de audio (50) según cualquiera de las reivindicaciones 1 a 6; una primera unidad de búfer (561) que almacena datos de audio que indican el habla convertida por la unidad de síntesis de habla e información de tiempo que indica el tiempo de una situación de la competición utilizada para generar el texto explicativo;

una segunda unidad de búfer (562) que almacena datos de vídeo que indican la situación de la competición en asociación con información de tiempo que indica la hora de la situación; y

una unidad de multiplexación (563) que extrae datos de audio en un momento designado de la primera unidad de búfer, el tiempo designado es una primera duración predeterminada antes de un tiempo actual, extrae datos de vídeo en el momento designado de la segunda unidad de búfer y multiplexa los datos de audio y los datos de vídeo.

8. El sistema de difusión (5) según la reivindicación 7,

en el que la unidad de gestión de mensajes (11) almacena los datos de la competición que incluyen el mensaje en asociación con información de tiempo que indica la hora de la situación de la competición indicada por el mensaje, en el que la unidad de generación de explicación (12) se refiere a los datos de competición en un momento como información no transmitida de la unidad de gestión de mensajes (11), siendo el momento una segunda duración predeterminada anterior al tiempo actual, y

en el que la segunda duración predeterminada es más corta que la primera duración predeterminada.

9. Un sistema de difusión (1, 5) que comprende:

el dispositivo de generación de orientación de audio (10, 30, 50) según cualquiera de las reivindicaciones 1 a 6, y un dispositivo transmisor (56) que transmite el texto explicativo y el habla.

10. Un procedimiento de generación de orientación de audio de un dispositivo de generación de orientación de audio (10, 30, 50), el procedimiento comprende:

una etapa de gestión de mensajes (S11, S12, S21, S23, S31, S32) de recepción y acumulación de un mensaje, en la que el mensaje se distribuye para cada actualización, en la que el mensaje son datos que representan una situación más reciente de una competición;

una etapa de generación de explicación (S41) de generación de un texto explicativo para transmitir información no transmitida detectada a partir del mensaje, en función de la información transmitida; y

una etapa de síntesis de habla (S46) de emisión de un habla convertida a partir del texto explicativo,

en la que la etapa de generación de explicación comprende una etapa (S47) de almacenamiento de la información no transmitida para el texto explicativo como la información transmitida después de generar el texto explicativo,

una etapa (S46) de espera hasta que finalice la emisión del habla, y

una etapa (S41) de inicio de un procedimiento para generar un nuevo texto explicativo en función de información no transmitida actualizada, **caracterizada por** las etapas de:

detectar una porción modificada del mensaje debido a la actualización del mensaje,
generar nueva información según un tipo de la parte variada y
acumular adicionalmente la nueva información generada.

FIG. 1

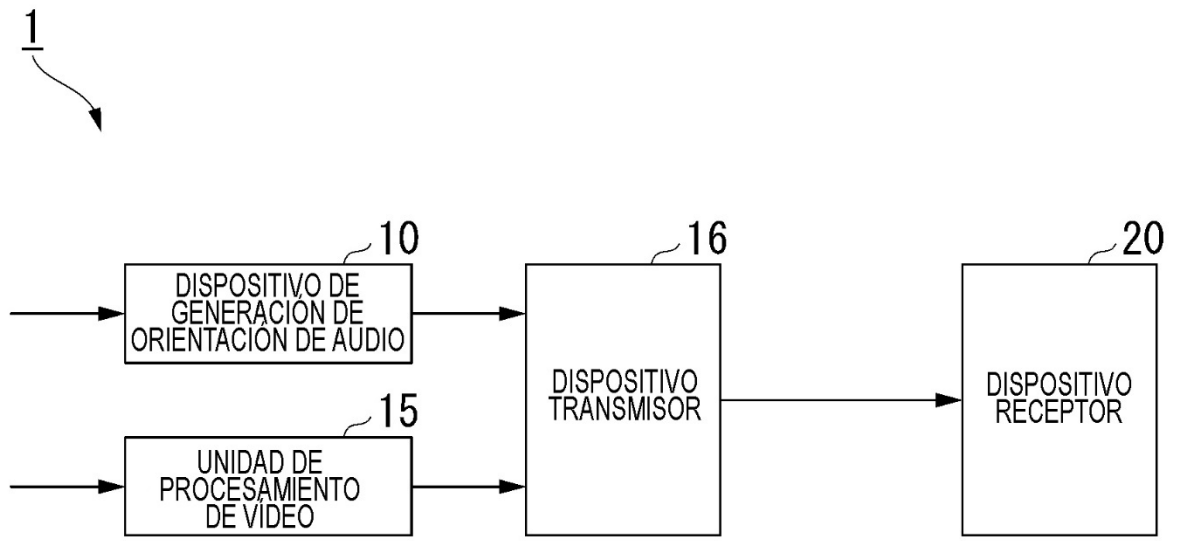


FIG. 2

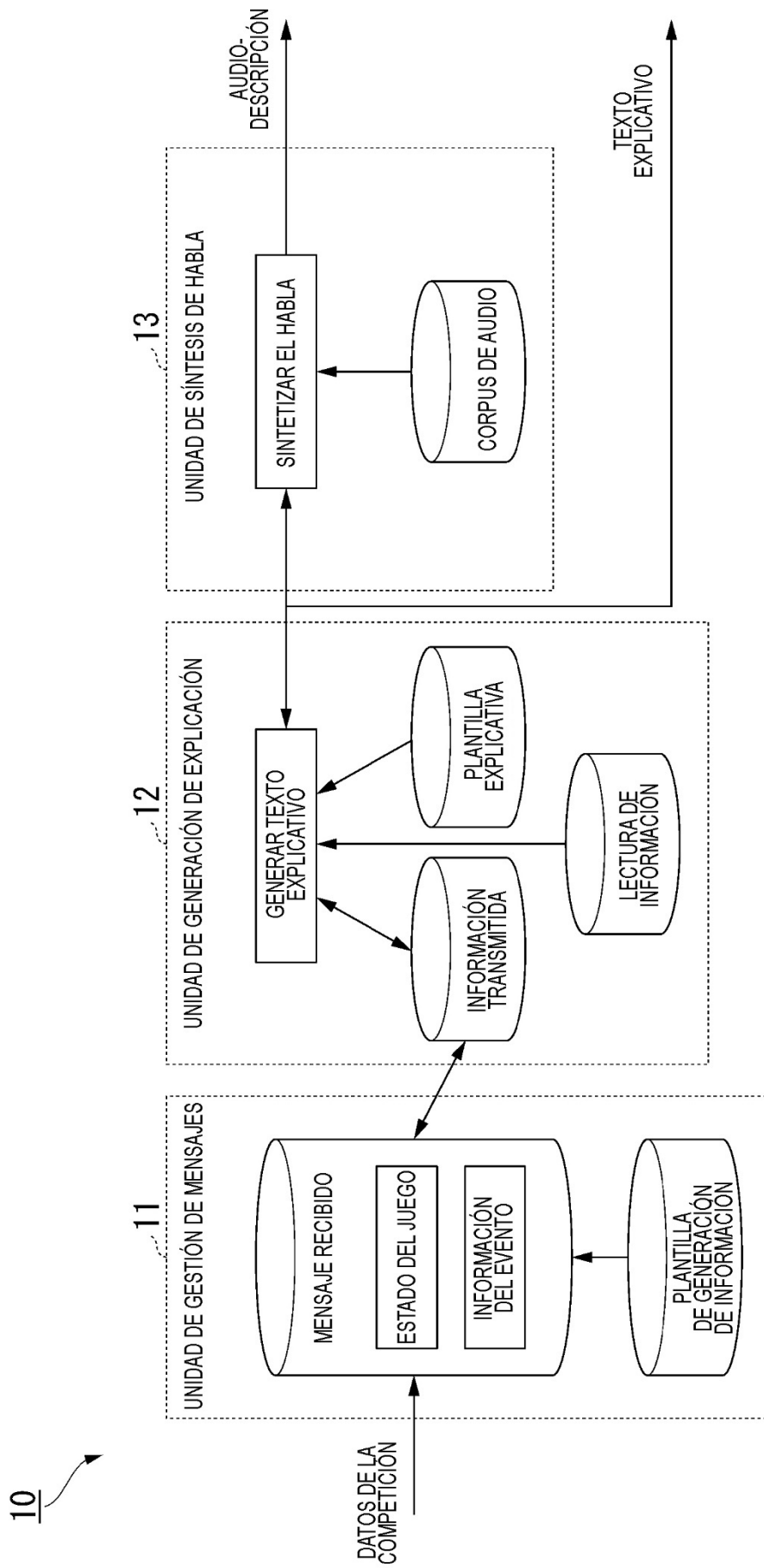


FIG. 3

```

TIPO DE MENSAJE: ESTADO DEL JUEGO
JUEGO: FINAL DE FÚTBOL MASCULINO
  ⋮
CARTA DE BATALLA [EQUIPO 1]: BRASIL
CARTA DE BATALLA [EQUIPO 2]: REPÚBLICA FEDERAL DE ALEMANIA
  ⋮
PUNTUACIÓN [EQUIPO 1]: 3
PUNTUACIÓN [EQUIPO 2]: 4,
  ⋮
    
```

FIG. 4

```

TIPO DE MENSAJE: LISTA DE EVENTOS
JUEGO: FINAL DE FÚTBOL MASCULINO
EVENTO [1]: {...},
EVENTO [2]: {...},
  ⋮
EVENTO [8]: {
  TIEMPO: PRIMERA MITAD 27 MINUTOS
  TIPO DE EVENTO: PATADA LIBRE (APUNTANDO DIRECTAMENTE A LA PORTERÍA)
  RESULTADO: ÉXITO
  ACTOR: {EQUIPO BRASIL, NOMBRE DEL JUGADOR: NM}
}
    
```

FIG. 5

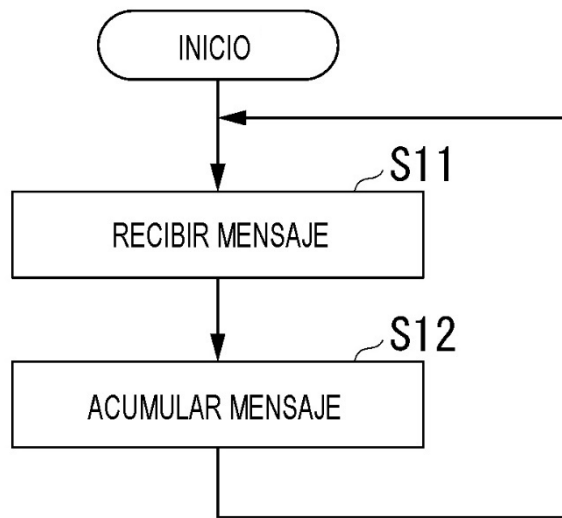


FIG. 6

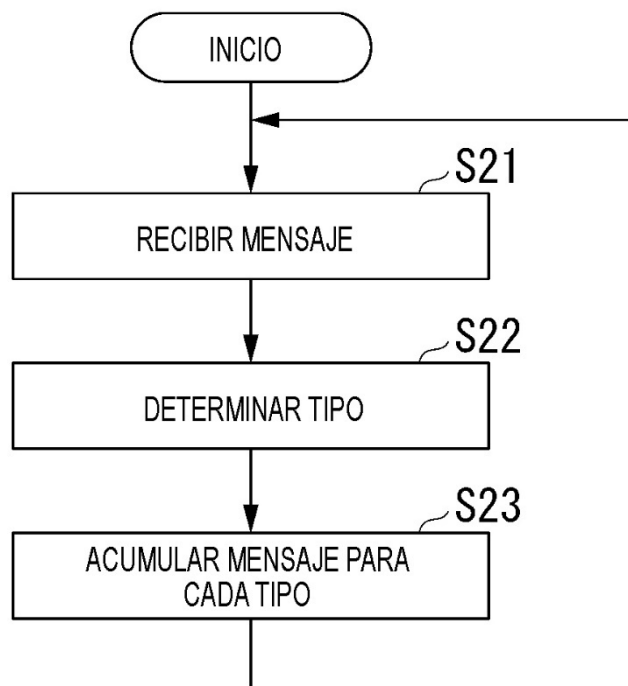


FIG. 7

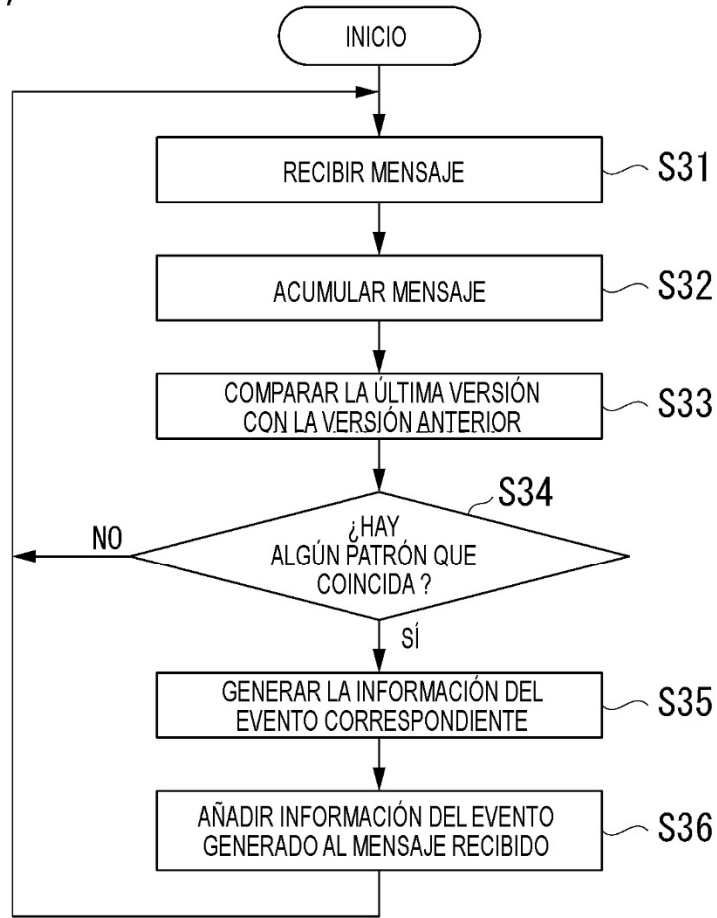


FIG. 8

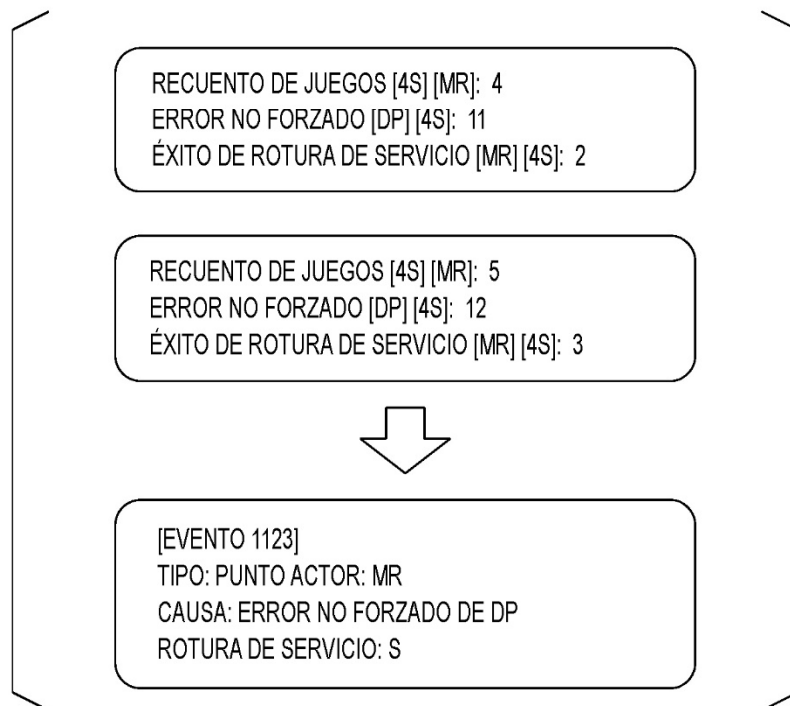


FIG. 9

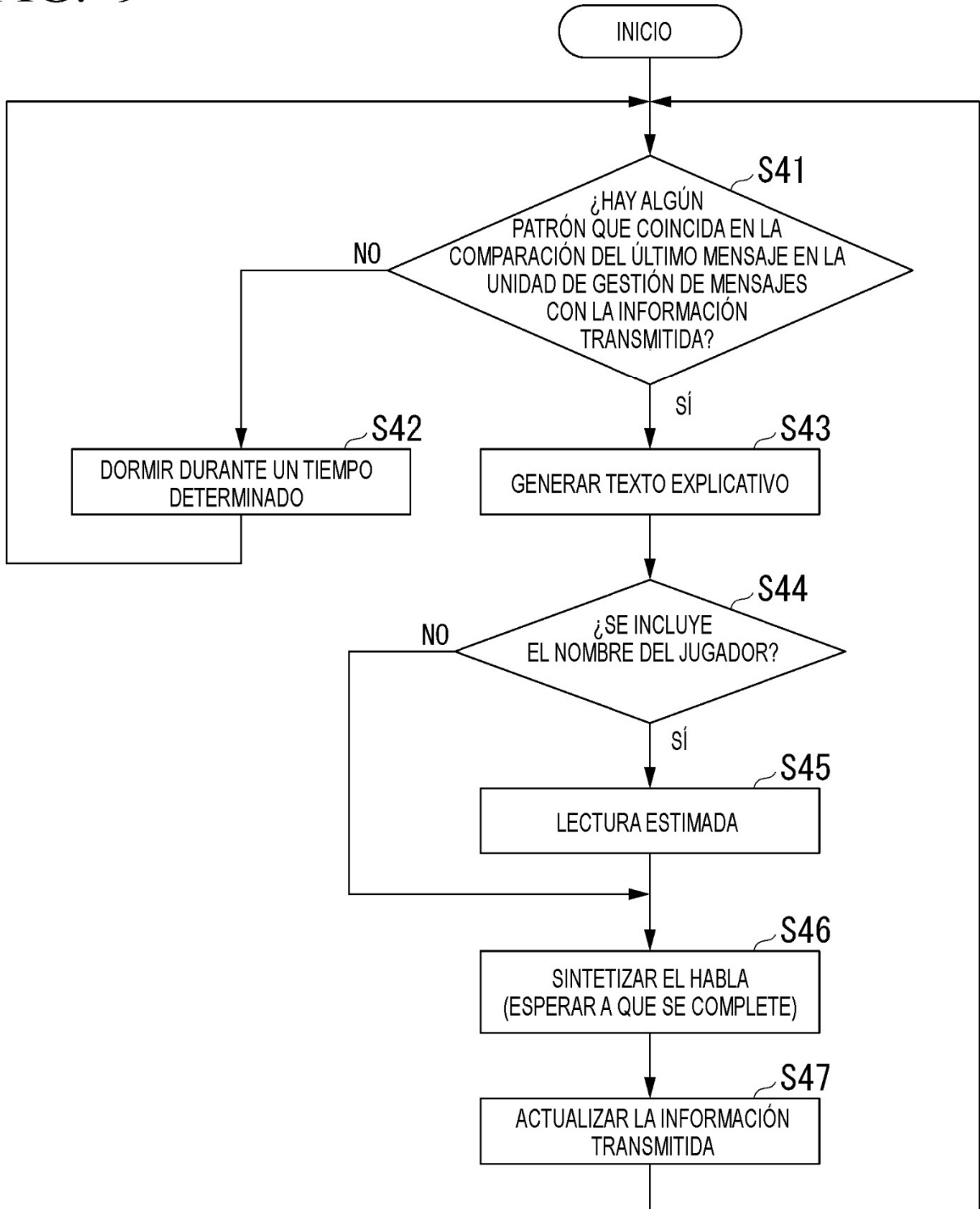


FIG. 10

PUNTO DE [NOMBRE COMPLETO DEL JUGADOR B]
[CAUSA]
[APELLIDO DEL JUGADOR B]. [NÚMERO TOTAL DE PARTIDAS]-ÉSIMO ÉXITO DE ROTURA DE SERVICIO
DEL JUEGO.
RECUENTO DE JUEGO, [APELLIDO DEL JUGADOR A] VS. [APELLIDO DEL JUGADOR B],
[NÚMERO DE JUEGOS ADQUIRIDOS POR EL JUGADOR A]-[NÚMERO DE ROTURAS DE SERVICIO DEL
JUGADOR B]-ÉSIMA ROTURA DE SERVICIO EN ESTE SET.

FIG. 11

TIPO DE EVENTO = PUNTO
[CAUSA] = ERROR NO FORZADO DEL JUGADOR A O DEL JUGADOR B
ROTURA DE SERVICIO = S

FIG. 12

PUNTO DE AD MR.
ERROR NO FORZADO DE DP.
MR. DÉCIMO ÉXITO DE ROTURA DE SERVICIO.
RECUENTO DE JUEGOS, DR VS. MR, 5-5.
TERCERA ROTURA DE SERVICIO EN ESTE SET.

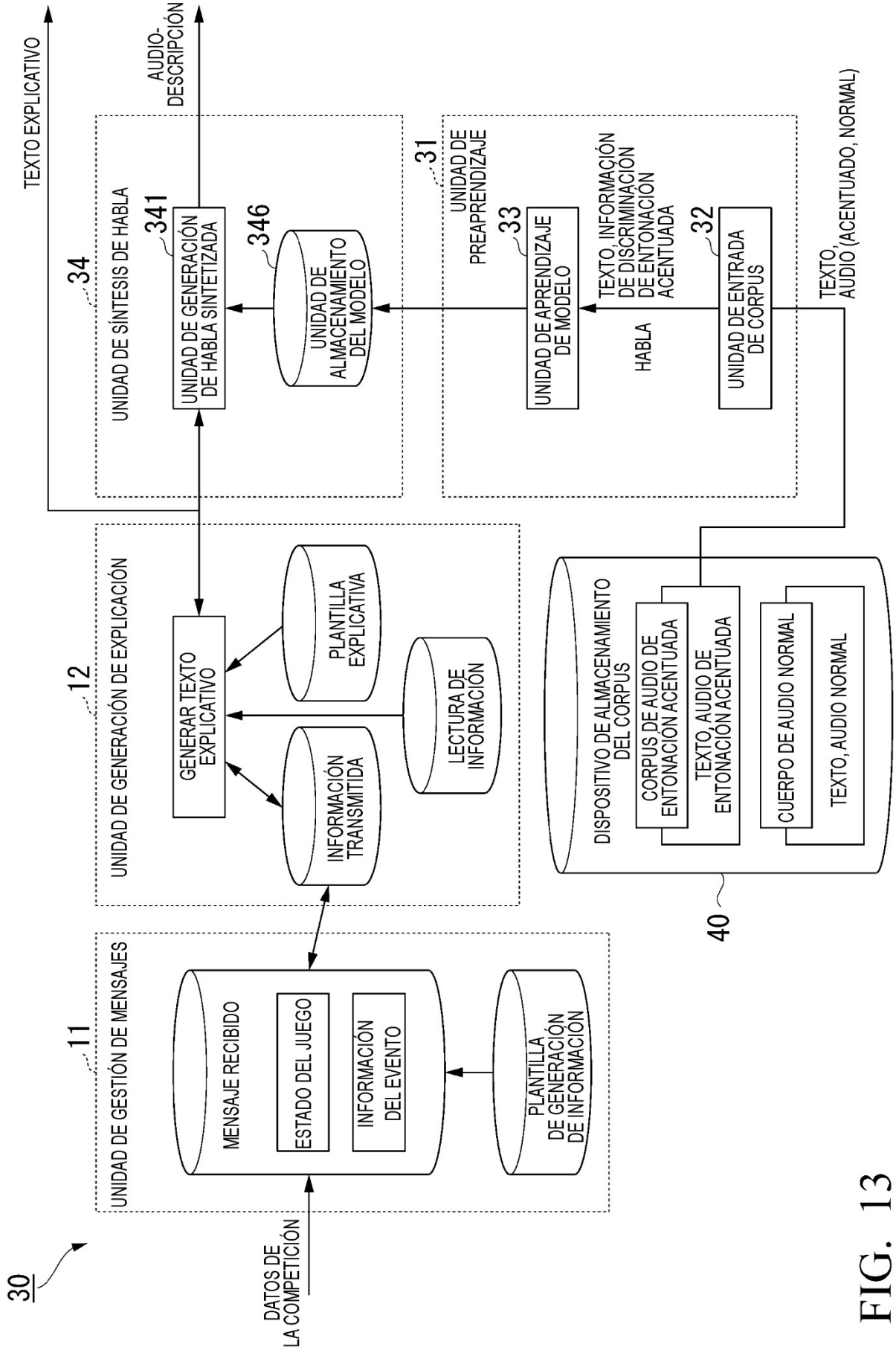


FIG. 13

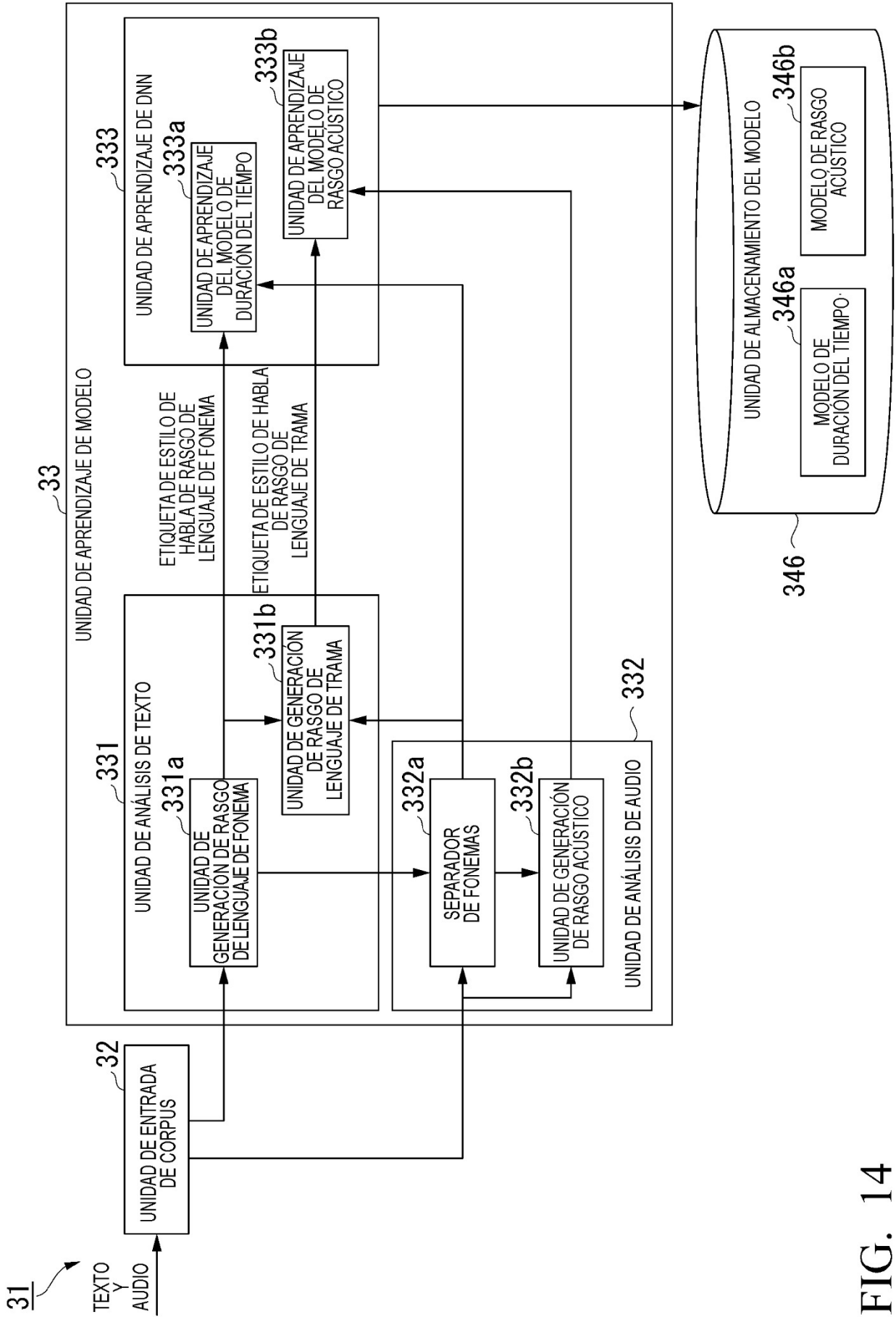


FIG. 14

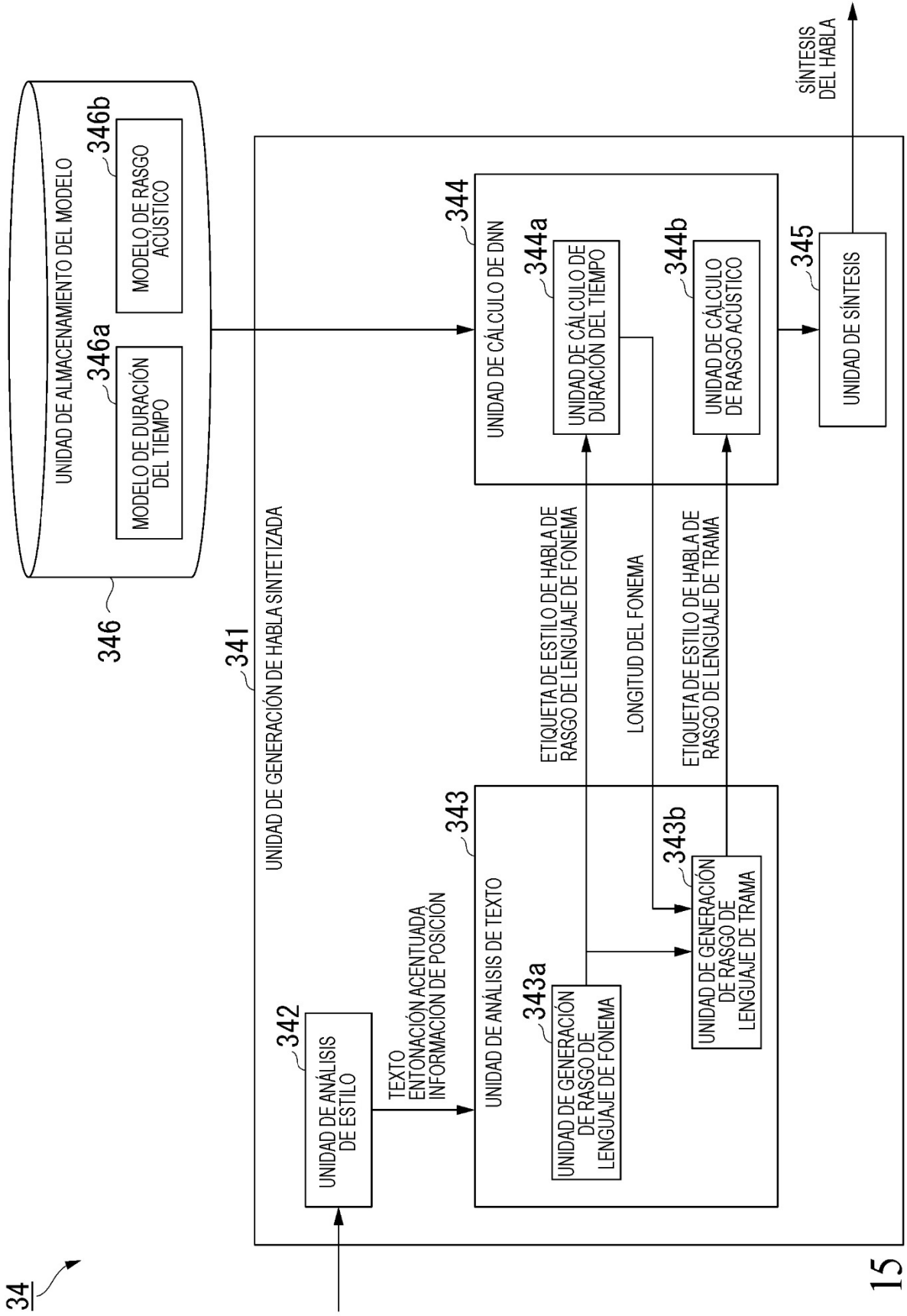


FIG. 15

FIG. 16

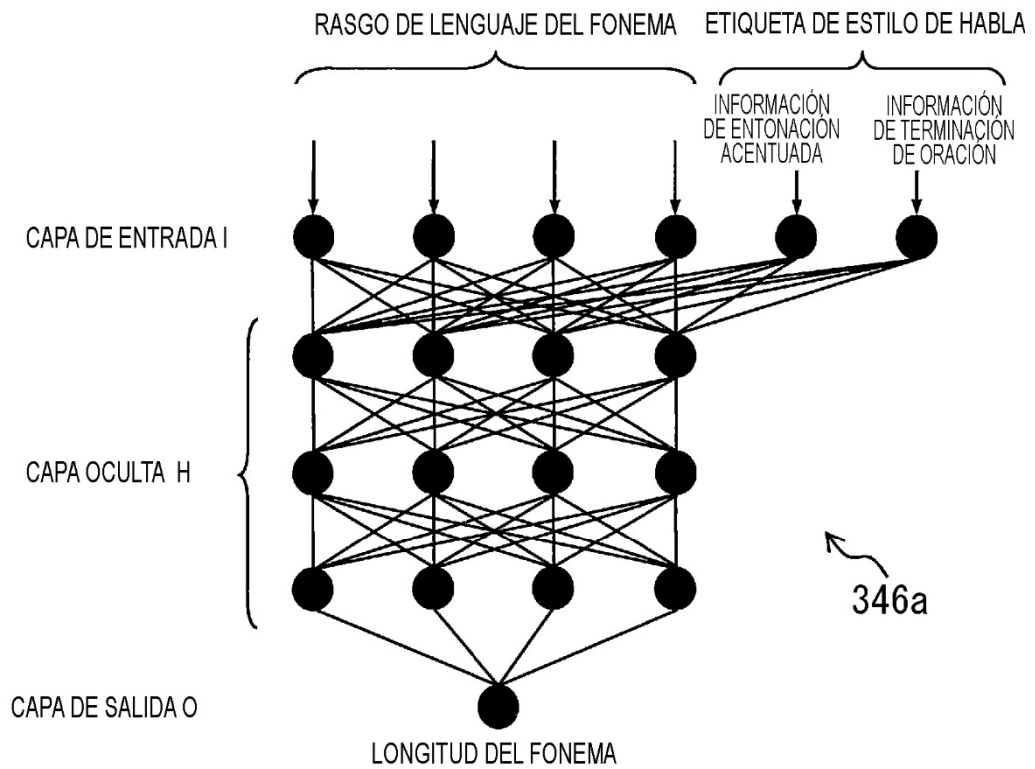


FIG. 17

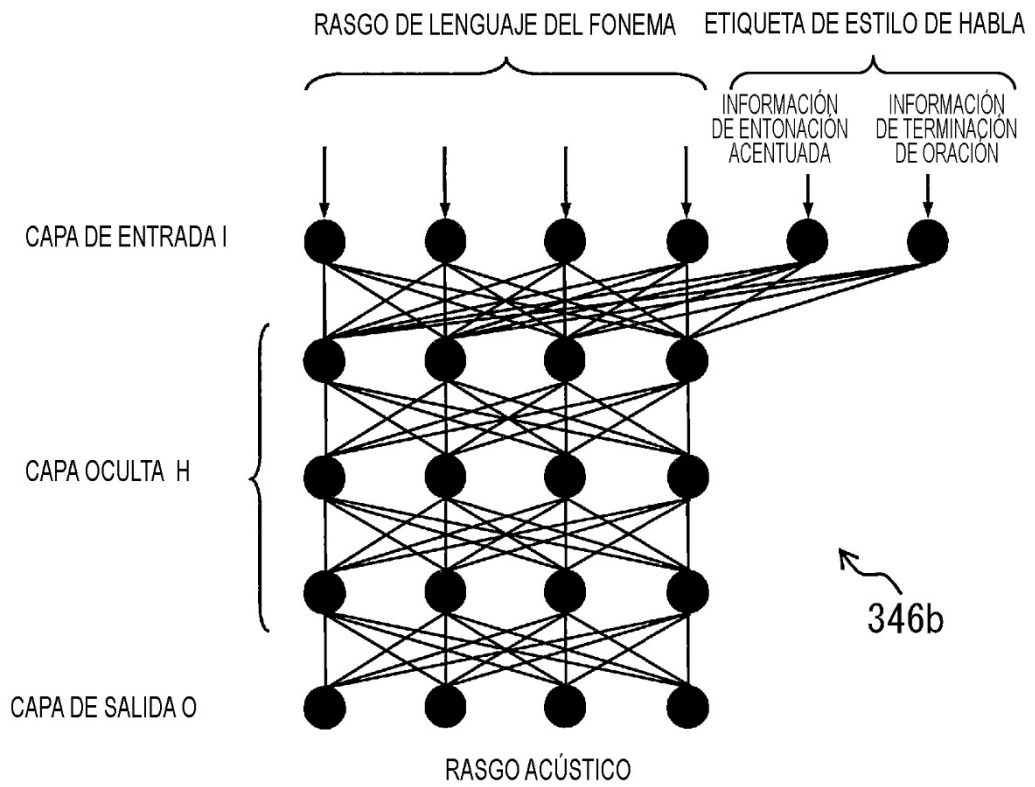


FIG. 18

JUGADORES QUE PARTICIPAN EN LA COMPETICIÓN
 [NOMBRE COMPLETO DEL JUGADOR A]= "A1 · A2"
 [NOMBRE COMPLETO DEL JUGADOR B]= "B1 · B2"
 [APELLIDO DEL JUGADOR A]= "A2"
 [APELLIDO DEL JUGADOR B]= "B2"
 ⋮

FIG. 19

LISTA DE EVENTOS
 EVENTO [1]
 ⋮
 EVENTO [123]
 [TIEMPO] = "○ HORA △ MINUTO □ SEGUNDO"
 [SET] = "4S"
 [RECUENTO DE JUEGOS] = "B2" , "5"
 [ERROR NO FORZADO] = "A2" , "12" .
 [ÉXITO DE ROTURA DE SERVICIO] = "B2" , "6"
 ⋮
 [NÚMERO TOTAL DE JUEGOS] = 10
 [NÚMERO DE JUEGOS ADQUIRIDOS POR EL JUGADOR A] = 5
 [NÚMERO DE JUEGOS ADQUIRIDOS POR EL JUGADOR B] = 5
 [NÚMERO DE ROTURAS DE SERVICIO DEL JUGADOR] = 6
 ⋮

FIG. 20

CONDICIONES: ERROR NO FORZADO DEL JUGADOR A, ÉXITO DE ROTURA DE SERVICIO DEL JUGADOR B
[NOMBRE COMPLETO DEL JUGADOR B] NO POINTO. [APELLIDO DEL JUGADOR A] NO ANFOOSUTOERAA. [APELLIDO DEL JUGADOR B], DAI [NÚMERO TOTAL DE JUEGOS] GEEMUBUREEKUSEEKOO. GEEMUKAUNTO, [APELLIDO DEL JUGADOR A] TAI [APELLIDO DEL JUGADOR B], [NÚMERO DE JUEGOS ADQUIRIDOS POR EL JUGADOR A] - [NÚMERO DE JUEGOS ADQUIRIDOS POR EL JUGADOR B]. KONOSETTO <ENTONACIÓN ACENTUADA> [NÚMERO DE ROTURAS DE SERVICIO DEL JUGADOR B] KAIME NO BUREEKU </ENTONACIÓN ACENTUADA>.

FIG. 21

CONDICIONES: ERROR NO FORZADO DEL JUGADOR A, ÉXITO DE ROTURA DE SERVICIO DEL JUGADOR B, NÚMERO DE ROTURAS DE SERVICIO DEL JUGADOR B \geq 5
[NOMBRE COMPLETO DEL JUGADOR B] NO POINTO. [APELLIDO DEL JUGADOR A] NO ANFOOSUTOERAA. [APELLIDO DEL JUGADOR B], DAI [NÚMERO TOTAL DE JUEGOS] GEEMUBUREEKUSEEKOO. GEEMUKAUNTO, [APELLIDO DEL JUGADOR A] TAI [APELLIDO DEL JUGADOR B], [NÚMERO DE JUEGOS ADQUIRIDOS POR EL JUGADOR A] - [NÚMERO DE JUEGOS ADQUIRIDOS POR EL JUGADOR B]. KONOSETTO <ENTONACIÓN ACENTUADA> [NÚMERO DE ROTURAS DE SERVICIO DEL JUGADOR B] KAIME NO BUREEKU </ENTONACIÓN ACENTUADA>.

FIG. 22

CONDICIONES: ERROR NO FORZADO DEL JUGADOR A, ÉXITO DE RUPTURA DEL JUGADOR B, NÚMERO DE ROTURAS DE SERVICIO DEL JUGADOR B $<$ 5
[NOMBRE COMPLETO DEL JUGADOR B] NO POINTO. [APELLIDO DEL JUGADOR A] NO ANFOOSUTOERAA. [APELLIDO DEL JUGADOR B], DAI [NÚMERO TOTAL DE JUEGOS] GEEMUBUREEKUSEEKOO. GEEMUKAUNTO, [APELLIDO DEL JUGADOR A] TAI [APELLIDO DEL JUGADOR B], [NÚMERO DE JUEGOS ADQUIRIDOS POR EL JUGADOR A] - [NÚMERO DE JUEGOS ADQUIRIDOS POR EL JUGADOR B]. KONOSETTO [NÚMERO DE ROTURAS DE SERVICIO DEL JUGADOR B] KAIME NO BUREEKU.

FIG. 23

B1 ▪ B2 NO POINTO.
A2. NO ANFOOSUTOERAA.
B2, DAI 10 GEEMUBUREEKU SEEK00.
GEEMUKAUNTO, A2 TAI B2, 5-5.
KONO SETTO <ENTONACIÓN ACENTUADA> 6 KAIME NO BUREEKU </ENTONACIÓN ACENTUADA>.

FIG. 24

B1 ▪ B2 NO POINTO.
A2. NO ANFOOSUTOERAA.
B2, DAI 10 GEEMUBUREEKU SEEK00.
GEEMUKAUNTO, A2 TAI B2, 5-5.
KONO SETTO 4 KAIME NO BUREEKU.

FIG. 25

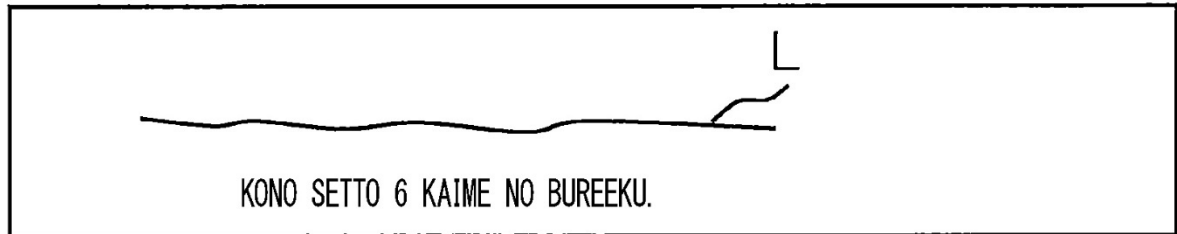


FIG. 26

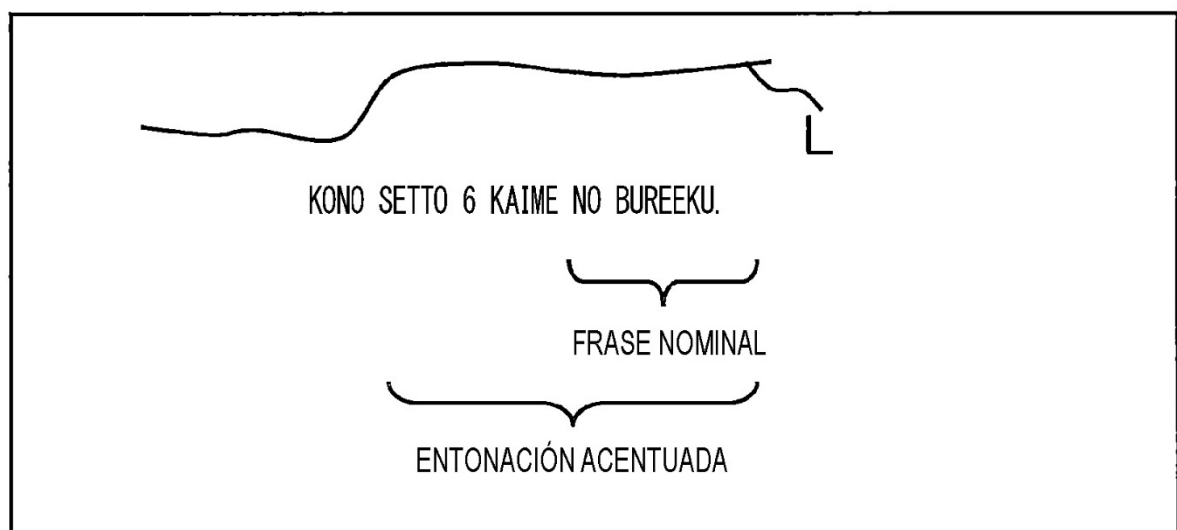


FIG. 27

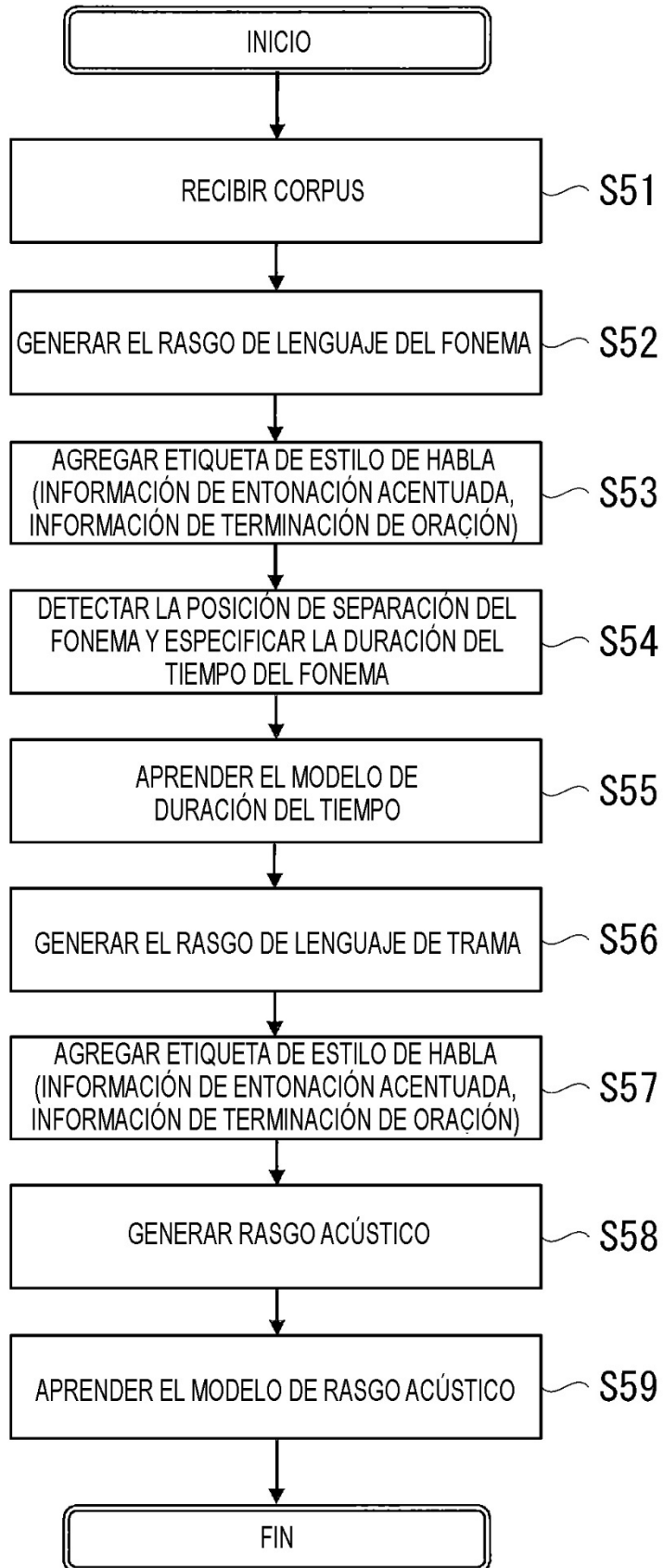


FIG. 28

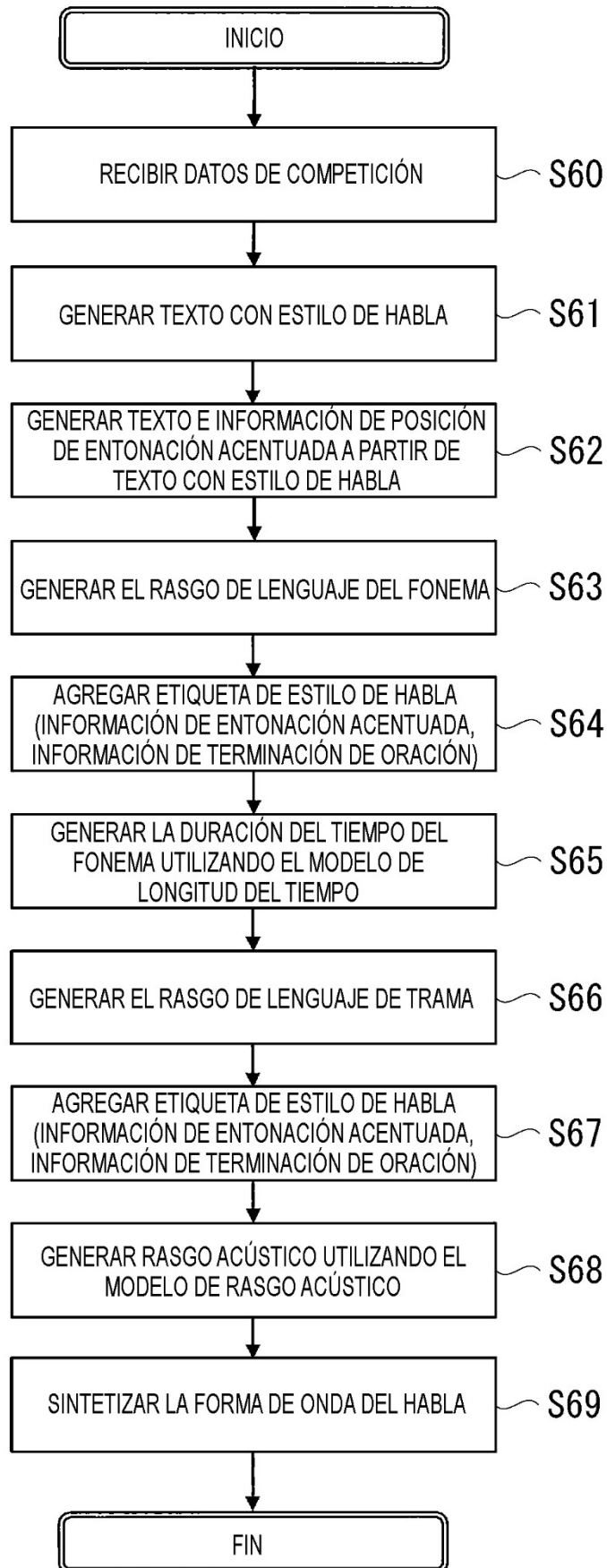


FIG. 29

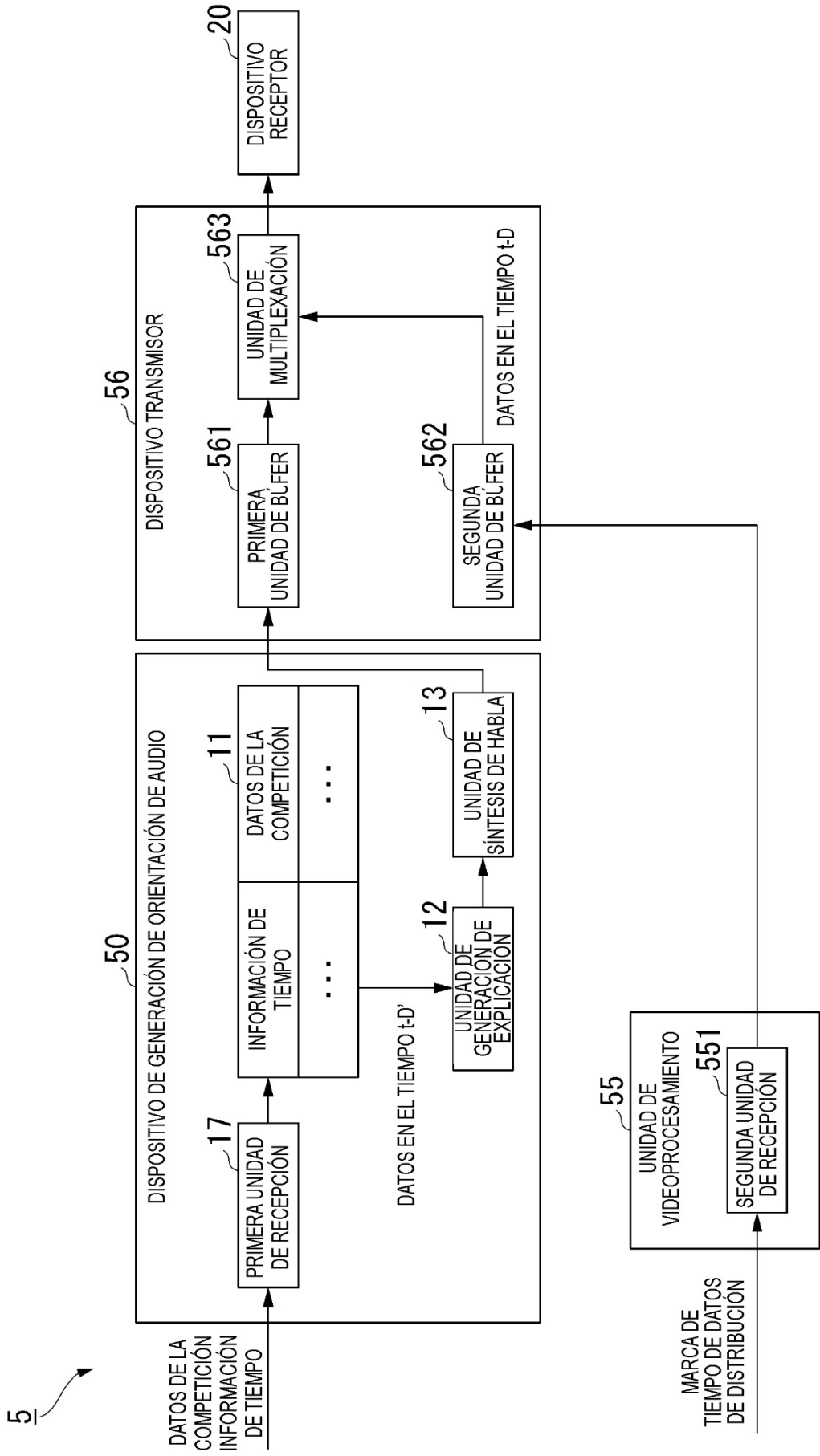


FIG. 30

