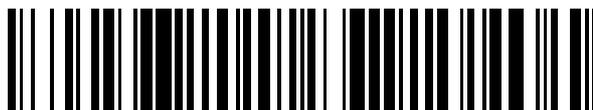


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 526 698**

51 Int. Cl.:

**G10L 15/22** (2006.01)

**H04M 3/493** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **06.07.2005 E 05768313 (8)**

97 Fecha y número de publicación de la concesión europea: **17.09.2014 EP 1779376**

54 Título: **Procedimiento, sistema y programa informático para recoger múltiples fragmentos de información durante un diálogo de usuario**

30 Prioridad:

**06.07.2004 US 885934**

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

**14.01.2015**

73 Titular/es:

**24/7 CUSTOMER, INC. (100.0%)  
910 E. Hamilton Avenue Suite 240  
Campbell, CA 95008, US**

72 Inventor/es:

**NGUYEN, PATRICK T. M.;  
LOPEZ-AMARO, JESUS;  
DESAI, AMIT V. y  
SHANA'A, ADEEB W. M.**

74 Agente/Representante:

**PONTI SALES, Adelaida**

ES 2 526 698 T3

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

## DESCRIPCIÓN

Procedimiento, sistema y programa informático para recoger múltiples fragmentos de información durante un diálogo de usuario

5

## ANTECEDENTES DE LA INVENCION

## Campo de la invención

10 **[0001]** La presente invención se refiere en general a sistemas de reconocimiento del habla. Más específicamente, se desvelan sistemas y procedimientos para construir una serie de interacciones con un usuario para recoger múltiples fragmentos de información relacionada con el fin de alcanzar una meta o materia específica (un diálogo multifranja) usando un enfoque basado en los componentes.

## 15 Descripción de la técnica relacionada

**[0002]** Los sistemas de reconocimiento del habla constituyen un procedimiento prometedor para automatizar funciones de servicio sin requerir cambios extensos en el comportamiento del usuario. Muchas empresas buscan ampliar o mejorar sus funciones de servicio al cliente usando tecnología de reconocimiento del habla para automatizar tareas que han sido manejadas tradicionalmente por agentes humanos. Para conseguirlo, los sistemas de reconocimiento del habla deben permitir al usuario preguntar y proporcionar información usando entradas habladas naturales en forma de conversación. Los recientes avances en ciertos ámbitos de la tecnología de reconocimiento del habla han ayudado a reducir algunos de los obstáculos tradicionales de los sistemas de reconocimiento del habla utilizables. Por ejemplo, los avances tecnológicos han permitido decodificar entradas habladas improvisadas según una amplia variedad de condiciones operativas realistas, como ruido de fondo y calidad imperfecta de la línea telefónica. Además, los avances recientes han permitido que las aplicaciones de voz reconozcan entradas de voz de una población de usuarios más amplia con diferentes acentos y estilos de habla.

**[0003]** Los sistemas de voz bien diseñados consiguen una alta aceptación del cliente. Por desgracia, la construcción de sistemas de voz eficaces usando enfoques del pasado ha sido difícil.

**[0004]** Los primeros enfoques requerían programación en las interfaces de programas de aplicación (API, *application program interfaces*) del motor de reconocimiento del habla. Estos enfoques cargaban a los desarrolladores con detalles específicos del motor de reconocimiento de bajo nivel tales como la manipulación de excepciones y la gestión de recursos. Por otra parte, como estas API eran específicas de un motor de reconocimiento en particular, las aplicaciones resultantes no podían ser transportadas fácilmente a otras plataformas.

**[0005]** La aparición de lenguajes de voz intermedios como VoiceXML como estándares abiertos simplificaron en cierta medida el proceso de desarrollo. Estos lenguajes de voz intermedios se acompañaron de una distribución de responsabilidades en un sistema de voz entre un navegador -que interpreta el lenguaje de voz y maneja la telefonía, el reconocimiento de voz y la infraestructura texto-habla- y una aplicación cliente -que proporciona el código de interacción de usuario (expresado en el lenguaje de voz). Como consecuencia, los desarrolladores de aplicaciones ya no tenían que preocuparse sobre las API de bajo nivel, aunque en su lugar eran responsables de generar documentos que serían ejecutados por el navegador de voz.

**[0006]** Aun con estos avances, sin embargo, el desarrollo de aplicaciones de voz seguía siendo complejo por diversos motivos. Por ejemplo, las aplicaciones de voz presentan un nuevo modelo de interacción con el usuario que es suficientemente distinto de la (bien conocida) interfaz gráfica de usuario como para necesitar experiencia de diseño e implementación especializada. Los conceptos de interfaz del habla, como la gestión de los diálogos, la optimización de la gramática y las interfaces multifranja, se implementan manualmente en todos los sistemas de voz preparados a medida. Dada la relativa novedad del paradigma del habla, esto impone cargas adicionales en los desarrolladores. Además, las demandas de aplicaciones para manejar la presentación, la lógica comercial y las funciones de acceso a los datos dieron como resultado arquitecturas fragmentadas que combinaban documentos generados de forma estática y dinámica, aplicaciones de servidor de segundo plano, gramáticas y otros componentes separados.

**[0007]** Se dispone de una serie de productos para simplificar el desarrollo de aplicaciones de voz para empresas. Un elemento central de muchos de estos productos es una biblioteca de componentes de voz

predefinidos y adaptables a medida cuyo uso reduce la cantidad de código que es preciso que desarrolle un programador. Estos componentes encapsulan habitualmente el código de lenguaje de voz, las gramáticas, los flujos de llamadas internas, los mensajes guía y las rutinas de recuperación de errores requeridos para obtener un fragmento de información del llamante, tal como la fecha, la hora, la cantidad de dólares, una secuencia de cifras o un elemento de un conjunto o una lista de elementos permitidos (por ejemplo, un conjunto de aeropuertos).

**[0008]** Una limitación importante de estos marcos de componentes es que los componentes no pueden combinarse para permitir que el usuario proporcione múltiples fragmentos de información en cada declaración. Por ejemplo, una aplicación de reservas de vuelos podría usar cuatro componentes: un aeropuerto de salida, un aeropuerto de destino, una fecha de salida y una hora de salida. Los marcos existentes permitirían a un usuario proporcionar los cuatro fragmentos de información en cuatro declaraciones separadas. Sin embargo, si la aplicación permitiera al usuario decir el aeropuerto de salida, el aeropuerto de destino y la fecha de salida en una declaración (por ejemplo "Mi vuelo es de Boston a San Francisco el lunes"), los componentes de aeropuerto de salida, aeropuerto de destino y fecha de salida no podrían combinarse de forma sencilla. En su lugar, se necesitaría desarrollar un nuevo componente con nuevas gramáticas, flujos de llamadas, mensajes guía, etc., para reconocer los dos aeropuertos y la fecha. Para llevar aún más lejos el ejemplo, si la aplicación permitiera al llamante conservar algunos fragmentos de información mientras cambia otros fragmentos de información (por ejemplo "No, en realidad vuelo a Oakland el martes"), habría que desarrollar un componente todavía más complejo.

**[0009]** Debido a estas limitaciones, las aplicaciones de voz que se basan en marcos de componentes existentes implementan diálogos altamente dirigidos en los que el flujo de llamadas está predeterminado en gran medida y cada etapa acepta sólo un único elemento de información, como en el intercambio ilustrado en la **FIG. 1a**. Dichos sistemas de voz son rígidos y a menudo penalizan al llamante que proporciona demasiada información, como en el intercambio ilustrado en la **FIG. 1b**. Como consecuencia, estos sistemas no son intuitivos ni eficaces, ya que no pueden capturar información rápidamente o adaptarse a las preferencias del usuario para proporcionar información.

**[0010]** Lo que se necesita es una aplicación de voz que use un enfoque más intuitivo, rápido y natural para obtener información de un usuario como, por ejemplo, un llamante.

**[0011]** El documento US-2002/0.038.213 describe un sistema y un procedimiento de procesamiento de un diálogo que comprende: una pluralidad de franjas, siendo cada franja una zona de almacenamiento que almacena un elemento de información predeterminado (por ejemplo, fecha de embarque, punto de salida, nombre de destino o vuelo) necesario para conseguir el propósito del diálogo con el usuario, por ejemplo, una reserva de billete aéreo; un analizador para extraer información correspondiente a los elementos de información de la información introducida por el usuario durante un diálogo y para almacenar la información extraída en la franja correspondiente; y un procesador de respuestas para producir una información de respuesta para un uso de acuerdo con el estado de almacenamiento de la información de las franjas.

**[0012]** El documento US-2003/013.089 describe un aparato de diálogo interactivo que genera mensajes guía y recibe respuestas, que procesa, y almacena los datos de las respuestas. El avance del diálogo se controla usando un modelo de estados que se analiza comparando las condiciones de transmisión con los datos de respuesta almacenados, deteniéndose en un estado para el que no se satisface la condición de bordes. Cada vez que se recibe una nueva respuesta el modelo se reanaliza desde el principio. Se generan mensajes guía como referencia a una historia de estados que se seleccionan buscando la correspondencia más larga entre las secuencias de historia y estado asociadas con mensajes guía almacenados. Los mensajes guía se comparan de manera que puede ponerse fin al diálogo en caso de una solicitud.

## RESUMEN DE LA INVENCION

**[0013]** En un aspecto, la presente invención proporciona un procedimiento según se expone en la reivindicación 1. En otro aspecto, la presente invención proporciona un sistema según se expone en la reivindicación 10. En otro aspecto, la presente invención proporciona un procedimiento según se expone en la reivindicación 17. En otro aspecto, la presente invención proporciona un producto de programa informático según se expone en la reivindicación 28.

**[0014]** Se desvelan sistemas y procedimientos para construir una serie de interacciones con un usuario con el fin de recoger múltiples fragmentos de información relacionada para alcanzar una meta o materia específica (un diálogo multifranja) usando un enfoque basado en componentes. Debe observarse que la presente invención puede

implementarse en numerosas formas, que incluyen un proceso, un aparato, un sistema, un dispositivo, un procedimiento o un medio legible por ordenador tal como un medio de almacenamiento legible por ordenador o una red informática en el que las instrucciones de programa se envían en líneas de comunicación ópticas o electrónicas. A continuación se describen varias realizaciones de la presente invención.

5

**[0015]** El procedimiento incluye en general la emisión de un mensaje guía de encabezamiento primario para producir los valores de las franjas en un segmento desde el usuario, la recepción de una respuesta de usuario primaria que contiene un valor para cada franja en al menos un subconjunto de las franjas en el segmento, el procesamiento de la respuesta de usuario primaria para determinar al menos un posible valor de reconocimiento para cada franja contenida en la respuesta de usuario primaria, el llenado de cada franja contenida en la respuesta de usuario primaria con un valor de correspondencia seleccionado entre los posibles valores de reconocimiento de correspondencia, y la repetición de la emisión, la recepción, el procesamiento y el llenado para cualquier franja no llena en el segmento hasta que se hayan llenado todas las franjas en el segmento de franjas.

10

**[0016]** El procedimiento puede incluir la realización de turnos para confirmar y/o aclarar los valores de franjas de correspondencia tales como aceptar en silencio una mejor correspondencia, confirmar pasivamente la mejor correspondencia, confirmar activamente la mejor correspondencia, desambiguar entre las mejores correspondencias y notificar al usuario el no reconocimiento. El procedimiento para confirmación y/o aclaración puede seleccionarse basándose, por ejemplo, en el número de posibles valores de reconocimiento para las franjas en la respuesta de usuario primaria y un nivel de confianza correspondiente para cada uno de los posibles valores de reconocimiento. Con una confirmación activa, se reconoce una respuesta de confirmación del usuario como una confirmación, una cancelación, o una cancelación y corrección. Con una cancelación y corrección, la corrección se procesa determinando al menos un posible valor de reconocimiento para cada franja contenida en la corrección. Con una confirmación pasiva, se emite un mensaje guía de confirmación pasiva con un siguiente mensaje guía de encabezamiento primario. El procedimiento puede incluir también la activación de las posibles franjas no llenas en el segmento de franjas, en el que el mensaje guía de encabezamiento primario produce valores para las franjas habilitadas en el segmento. El procedimiento puede basarse en texto o en el habla.

15

20

25

**[0017]** Estas y otras características y ventajas de la presente invención se presentarán con más detalle en la siguiente descripción detallada y en las figuras adjuntas que ilustran a modo de ejemplo los principios de la invención.

30

#### BREVE DESCRIPCIÓN DE LOS DIBUJOS

**[0018]** La presente invención se comprenderá más fácilmente mediante la siguiente descripción detallada conjuntamente con los dibujos adjuntos, en los que números de referencia iguales designan los mismos elementos estructurales.

35

Las **FIG. 1a** y **FIG. 1b** ilustran ejemplos de diálogos altamente dirigidos usando aplicaciones de voz convencionales.

40

Las **FIG. 2a-2c** ilustra varios ejemplos de diálogos multifranja.

La **FIG. 3** es un diagrama de bloques de un sistema de aplicación de voz multifranja ilustrativo.

La **FIG. 4** ilustra una realización de un marco para gestionar una conversación basada en reconocimiento del habla multifranja.

45

La **FIG. 5** ilustra ejemplos de segmentos para un diálogo de reconfirmación de vuelo multifranja.

La **FIG. 6** ilustra el contenido de uno de los segmentos de ejemplo, es decir, el itinerario de vuelo, en el diálogo de reconfirmación de vuelo multifranja de la **FIG. 5**.

50

La **FIG. 7** ilustra ejemplos de intercambios que intervienen en un diálogo de itinerario de vuelo.

La **FIG. 8** es un organigrama que ilustra varias etapas de un diálogo multifranja de ejemplo.

55

La **FIG. 9** es un organigrama de un intercambio de ejemplo que usa confirmación activa.

La **FIG. 10** es un organigrama de un intercambio de ejemplo que usa confirmación pasiva.

La **FIG. 11** es un organigrama que ilustra un ejemplo de una conversación que usa una funcionalidad de retroceso de un sistema de diálogo multifranja.

5 La **FIG. 12** es un organigrama que ilustra un ejemplo de una conversación que usa una funcionalidad de cambio del sistema de diálogo multifranja.

La **FIG. 13** es un organigrama que ilustra un ejemplo de una conversación que usa una funcionalidad de revisión del sistema de diálogo multifranja.

10

Las **FIG. 14** y **15** ilustran algunos de los posibles objetos de grupos de franjas para las clases de grupos de franjas de reconocimiento anterior y reconocimiento posterior para un segmento de itinerario de vuelo, respectivamente.

La **FIG. 16** es un diagrama de bloques que ilustra una estructura de flujo de diálogo de ejemplo.

15

Las **FIG. 17** y **18** son organigramas que ilustran un procesamiento de ejemplo de la entrada del usuario en modos de confirmación normal y pasivo, respectivamente.

La **FIG. 19** ilustra una posible gramática para un intercambio con tres franjas.

20

La **FIG. 20** ilustra una posible gramática para un intercambio de confirmación que incluye gramáticas de aceptación y cancelación así como gramáticas para las tres franjas que se confirman.

## DESCRIPCIÓN DE REALIZACIONES ESPECÍFICAS

25

**[0019]** Se desvelan sistemas y procedimientos para construir una serie de interacciones con un usuario con el fin de recoger múltiples fragmentos de información relacionada para conseguir una meta o materia específica (un diálogo multifranja) usando un enfoque basado en componentes. Los sistemas y procedimientos de diálogo multifranja obtienen información de un usuario realizando una serie de interacciones basadas en reconocimiento del habla. Los sistemas y procedimientos incluyen la determinación de los mensajes guía emitidos hacia el usuario así como las gramáticas y las reglas semánticas usadas para reconocer las entradas de usuario tales como declaraciones en cada punto de la conversación o intercambio. La siguiente descripción se presenta para facilitar que cualquier experto en la materia realice y haga uso de la invención. Las descripciones de realizaciones y aplicaciones específicas se proporcionan sólo como ejemplos y las diversas modificaciones serán evidentes para los expertos en la materia. Los principios generales definidos en la presente memoria descriptiva pueden aplicarse a otras realizaciones y aplicaciones sin apartarse del espíritu y el alcance de la invención. Así, la presente invención debe adecuarse al máximo alcance que comprende numerosas alternativas, modificaciones y equivalentes consistentes con los principios y las características desvelados en la presente memoria descriptiva. Para mayor claridad, los detalles relativos al material técnico que se conoce en los campos técnicos relacionados con la invención no se han descrito en detalle para que no confundan innecesariamente la presente invención.

30

35

40

**[0020]** Las **FIG. 2a-2c** ilustran varios ejemplos de diálogos multifranja. Un diálogo multifranja tiene el objetivo de recoger múltiples fragmentos de información relacionada ("franjas") para conseguir una meta o materia específica, tal como la localización de una reserva aérea. Con el fin de conseguir una interacción de usuario con éxito con la comodidad y la facilidad esperadas por los seres humanos, una aplicación de diálogo multifranja maneja preferentemente ciertos comportamientos e interacciones típicos de las interacciones humanas en un medio hablado, lo que incluye que: (i) el llamante puede proporcionar las franjas en un orden arbitrario, (ii) el llamante puede proporcionar múltiples franjas en una sola entrada como, por ejemplo, una declaración hablada, (iii) el llamante puede proporcionar sólo un subconjunto de franjas solicitadas por la aplicación en una sola declaración, (iv) el llamante puede aclarar o corregir la interpretación por la aplicación de las franjas que ha proporcionado el llamante, (v) el llamante puede modificar las franjas previas en declaraciones posteriores.

45

50

**[0021]** Para satisfacer estos requisitos de la interacción humana, una aplicación de diálogo puede realizar una ruta de llamada larga y sofisticada de considerable complejidad. Sin embargo, las aplicaciones de voz convencionales no están bien adaptadas para implementar el diálogo multifranja. En particular, el orden dinámico y las combinaciones en que puede proporcionarse la información no son manejadas fácilmente por los marcos existentes de componentes y las metodologías de desarrollo de aplicaciones de voz convencionales que especifican flujos de llamadas rígidos y predeterminados.

55

**[0022]** En lugar del enfoque manual, los sistemas y procedimientos de diálogo multifranja tal como se describen en la presente memoria descriptiva pueden usarse para construir aplicaciones de diálogo multifranja usando un enfoque basado en componentes. Dicho enfoque basado en componentes automatiza el proceso de elaboración de multifranjas con componentes para los elementos conductuales de un intercambio de múltiples franjas entre el ser humano y la máquina, lo que incluye estrategias de secuenciación, agrupación, manejo de mensajes guía, confirmación y/o modificación.

**[0023]** La **FIG. 3** es un diagrama de bloques de un sistema de aplicación de voz ilustrativo 300 que incluye generalmente un teléfono 301 en comunicación con un navegador de voz 303 por medio de una red telefónica 302 que a su vez está en comunicación con una aplicación de voz 309 por medio de una red de datos 308. El navegador de voz 303 incluye el hardware y el software para realizar comunicaciones de audio bidireccionales con un llamante por medio de la red telefónica 302 y el teléfono 301. El navegador de voz 303 puede ejecutar un programa expresado en un lenguaje de voz transmitido en forma de documentos por la red de datos 308 (como Internet o una intranet) desde la aplicación de voz 309. El navegador de voz 303 y la aplicación de voz 309 pueden residir en cualquiera de los diversos sistema informáticos disponibles.

**[0024]** El lenguaje de voz puede ser un lenguaje de marcadores como VoiceXML o Speech Application Language Tags (SALT). El navegador de voz 303 puede incluir diversos componentes tales como un motor de reconocimiento del habla 304, un sintetizador de texto-habla 305, un aparato de reproducción de audio 306 para reproducir de forma audible archivos registrados usando formatos de audio disponibles en general y un componente para tratamiento de llamadas en la red telefónica 307. Para construir aplicaciones de voz pueden emplearse componentes disponibles comercialmente y/o componentes de propiedad exclusiva en la implementación de alguno o todos los diversos componentes del navegador de voz 303.

**[0025]** El navegador de voz 303 puede ser responsable de la detección de una llamada entrante, la contestación de la llamada entrante, la petición de un documento de voz inicial de la aplicación de voz 309 y la interpretación del documento de voz y la ejecución de las instrucciones contenidas en el documento de voz según las reglas del lenguaje de voz aplicable. Las instrucciones pueden incluir la emisión de mensajes guía audibles hacia el usuario y el procesamiento de respuestas de voz desde el usuario usando gramáticas especificadas. En la producción de los mensajes guía para el usuario, el navegador de voz 303 puede usar la reproducción de un reproductor de audio 306 para reproducir de forma audible mensajes registrados previamente o puede usar el sintetizador de texto-habla 305 para reproducir de forma audible mensajes de texto. El navegador de voz 303 puede generar a continuación acciones en respuesta para los eventos de usuario (por ejemplo, entrada hablada o desconexión de llamada) o eventos de sistema (por ejemplo, excepciones de subsistemas), y/o acciones que pueden ser definidas en el documento actual o en otro documento que se tomará desde la aplicación de voz 309.

**[0026]** Una vez descrita la arquitectura general del sistema de aplicación de voz 300, a continuación se presentará un marco multifranja implementado por el sistema de aplicación de voz 300, como, por ejemplo, una combinación del navegador de voz 303 y aplicación de voz 309. Un sistema que implementa el modo de conversación multifranja solicita generalmente al usuario preguntas más abiertas mientras que un sistema que implementa un modo de conversación convencional de diálogo dirigido solicita al usuario con mensajes guía más específicos y acepta un conjunto de respuestas más limitado. La **FIG. 4** ilustra una realización de un marco multifranja 401 para gestionar una conversación basada en reconocimiento del habla multifranja. En particular, el marco multifranja 401 implementa un modo multifranja para obtener información de un usuario. El marco 401 puede incluir una aplicación multifranja 402 que gestiona la lógica del negocio y las responsabilidades de acceso a los datos para un conjunto de franjas 403. En la presente memoria descriptiva una franja se refiere como un elemento de datos cuyo valor puede obtenerse a partir de una entrada de usuario tal como una entrada de texto o una declaración hablada. Por ejemplo, una aplicación de reservas de vuelos puede gestionar una franja de aeropuerto de salida cuyos valores válidos se encuentran en una base de datos de planificación de vuelos, y cuyo valor, una vez pronunciado o introducido de otra forma por el usuario, debería almacenarse en un registro de reservas. Una plataforma multifranja 404 construye una o más interacciones de voz para obtener las franjas deseadas desde el usuario.

**[0027]** La **FIG. 5** ilustra ejemplos de segmentos para un diálogo de reconfirmación de vuelo multifranja y la **FIG. 6** ilustra el contenido de uno de los segmentos de ejemplo, es decir, el itinerario de vuelo, en el diálogo de reconfirmación de vuelo multifranja. En particular, un diálogo multifranja puede descomponerse en subunidades en varios niveles. En el nivel inferior, un turno es un flujo ininterrumpido de entradas (por ejemplo, habla) de un participante, es decir, el sistema o el usuario. Así un turno de sistema es aquel en el que el sistema solicita al usuario un mensaje y un turno de usuario es aquel en el que el usuario hace una declaración que el sistema intenta

interpretar. Un intercambio es una serie de turnos que captura los valores para una o más franjas desde el usuario. Así un intercambio puede incluir uno o más turnos de confirmación, corrección o desambiguación hasta que el sistema acepta un conjunto único y final de valores de franjas. Un segmento es una serie de intercambios que captura un conjunto de franjas relacionadas. El segmento es la unidad más grande de un diálogo que permite proporcionar todas sus franjas asociadas en un turno de usuario. En el ejemplo mostrado en la **FIG. 6**, las cinco franjas, es decir, aeropuerto de salida, aeropuerto de destino, fecha, hora e indicador mañana/tarde, forman el segmento de itinerario de vuelo de manera que hasta las cinco franjas del segmento de itinerario de vuelo pueden introducirse en un turno de usuario, por ejemplo por el usuario que pronuncia “de San Francisco a Boston mañana a las 9 de la mañana”.

10

**[0028]** El diálogo multifranja puede variar en complejidad desde diálogos de un solo segmento a diálogos complejos multisegmento en los que cada segmento incluye una o más franjas. Un ejemplo de un diálogo de un solo segmento es una aplicación de atención telefónica que solicita el nombre de empleado al llamante. En cambio, un ejemplo de un diálogo multisegmento es un diálogo de reconfirmación de vuelo como el mostrado en la **FIG. 5**. Por ejemplo, el diálogo de reconfirmación de vuelo puede solicitar un itinerario de vuelo al llamante (que puede incluir franjas para los aeropuertos de salida y de destino así como la fecha, la hora y el indicador de mañana/tarde de la salida tal como se muestra en la **FIG. 6**), y puede verificar la identidad del llamante solicitando un localizador de registro y solicitando detalles personales, como el apellido.

15

**[0029]** La **FIG. 7** ilustra ejemplos de intercambios que intervienen en un diálogo de itinerario de vuelo. El segmento de itinerario de vuelo 701 contiene 5 franjas que representan un itinerario de vuelo. En el primer intercambio 702, el sistema puede preguntar al usuario “¿Cuál es su itinerario de vuelo?” para permitir que el usuario introduzca información o valores para hasta las 5 franjas. El usuario puede responder con “Volaré desde Boston” y después de cualquier turno de confirmación/aclaración aplicable, la aplicación acepta Boston como valor para la franja del aeropuerto de salida. Según la lógica predefinida, la aplicación determina, por ejemplo, que sólo se solicitará el aeropuerto de destino en el siguiente intercambio 703, y emite un mensaje guía “¿Y cuál es su destino?”. Después de que se procesa la respuesta del usuario, se almacena el valor San Francisco en la franja del aeropuerto de destino. La aplicación puede solicitar a continuación las tres franjas no llenas restantes en el siguiente intercambio 704, por ejemplo, emitiendo un mensaje guía “¿Y cuándo va a volar?”. La respuesta del usuario de “el viernes que viene a las nueve y media de la mañana” llena las tres franjas que quedan y el diálogo se completa.

20

**[0030]** La **FIG. 8** es un organigrama que ilustra una conversación multifranja de ejemplo 800. Al principio del intercambio en el bloque 802, el sistema de aplicación de voz emite un mensaje guía de encabezamiento primario para producir valores para un conjunto de franjas del usuario. El mensaje guía del sistema se pronuncia normalmente de manera que el usuario puede proporcionar potencialmente valores para una serie de franjas deseadas en un turno. En el bloque 804, el usuario responde con un turno de usuario primario en respuesta a la emisión del mensaje guía. El turno de usuario puede proporcionar uno o más valores de franjas. Si el turno de usuario es reconocido por el sistema, puede seguir una serie de turnos de confirmación/aclaración en el bloque 806 en los que el sistema puede aceptar opcionalmente en silencio la mejor correspondencia, confirmar pasivamente la mejor correspondencia, solicitar la confirmación activa de la mejor correspondencia, desambiguar entre las mejores correspondencias o notificar al usuario un no reconocimiento, por ejemplo, tal como se describirá en más detalle más adelante. Cuando los valores de franjas del turno son confirmados de forma explícita o implícita, las franjas se llenan con sus valores respectivos en el bloque 808.

25

**[0031]** Los bloques 802-808 representan un intercambio. Cuando se completa el intercambio actual, el sistema determina si quedan franjas no llenas en el segmento actual en el bloque de decisión 810. Si todas las franjas en el segmento actual están llenas, el sistema determina si existen segmentos adicionales de franjas en el bloque de decisión 812. Si todas las franjas en todos los segmentos están llenas, entonces el diálogo multifranja 800 se ha completado. Alternativamente, si quedan franjas no llenas en el segmento actual y/o si quedan segmentos de franjas adicionales, el diálogo 800 regresa al bloque 802 para iniciar el siguiente intercambio. Por ejemplo, el siguiente intercambio puede seguir con el intercambio anterior manejando mensajes guía para cualquier franja que no se hubiera llenado (por ejemplo, “... ¿y aproximadamente a qué hora?”) o continuar en el siguiente segmento de franjas gestionando los mensajes guía para un nuevo conjunto de franjas (por ejemplo, “¿Qué tipo de vehículo desea alquilar?”).

30

**[0032]** Tal como se observa anteriormente, si el turno de usuario es reconocido, puede seguir una serie de turnos de confirmación y/o aclaración en la que el sistema puede solicitar opcionalmente confirmación activa de la mejor correspondencia, confirmar pasivamente la mejor correspondencia, desambiguar entre las mejores correspondencias, aceptar en silencio la mejor correspondencia o notificar al usuario un no reconocimiento, por

35

40

45

50

55

ejemplo. A continuación se describirán los turnos de confirmación/aclaración en más detalle con referencia a las **FIG. 9 y 10**.

**[0033]** En particular, la **FIG. 9** es un organigrama de un intercambio de ejemplo que usa confirmación activa.

5 En una confirmación activa, se emite un mensaje guía de confirmación para el usuario (por ejemplo, "Creo que ha dicho de Austin a San Francisco. ¿Es correcto?"). El mensaje guía de confirmación puede prolongarse durante un intercambio si existen varios ciclos de confirmación. El sistema puede requerir que el usuario acepte explícitamente el o los valores para las franjas asociadas antes de llenar las franjas. El usuario puede cancelar los valores reconocidos anteriores diciendo una frase de cancelación como "No" o "No, no es correcto". Además, para facilitar  
10 una interacción más eficaz, el sistema puede aceptar opcionalmente una declaración del usuario que incluya una frase de cancelación seguida por una corrección como "No, he dicho de Boston a San Francisco". Si el usuario cancela los valores reconocidos anteriores, el sistema puede borrar todos los valores de franjas reconocidos en el turno de usuario y reproducir un mensaje como "Inténtelo otra vez. ¿Cuál es su itinerario de vuelo?" tal como se muestra en la **FIG. 9** e iniciar de nuevo el intercambio. Alternativamente, el sistema puede tratar la respuesta del  
15 usuario pronunciada en el turno de usuario como una corrección y repetir el turno de confirmación/aclaración, por ejemplo, indicando "Perfecto. De Boston a San Diego. ¿Es correcto?" Si la corrección omite algunas de las franjas que se están confirmando, el sistema puede conservar los valores pronunciados anteriormente de dichas franjas.

**[0034]** La **FIG. 10** es un organigrama de un intercambio de ejemplo que usa confirmación pasiva. En una

20 confirmación pasiva, el sistema de diálogo multifranja emite un mensaje guía que es una combinación de un mensaje guía de confirmación pasiva y un mensaje guía de encabezamiento para el siguiente intercambio. A modo de ejemplo, el mensaje guía producido combinado puede ser, por ejemplo, "Perfecto. De Austin a San Francisco. ¿Y cuándo va a volar?". Como no es necesario que el usuario pronuncie explícitamente una confirmación si los valores son correctos, la técnica de confirmación pasiva facilita la velocidad del diálogo y así es adecuada para confirmar los  
25 resultados del reconocimiento cuando el sistema tiene un alto nivel de confianza. Cuando se usa confirmación pasiva, las franjas se llenan con sus valores respectivos y esos valores llenados pueden eliminarse o recuperarse cuando el usuario pronuncia una frase de cancelación tal como "No" o "No, no es correcto", o una frase de cancelación seguida por una corrección como "No es correcto. He dicho de Boston a San Francisco". Si el usuario transmite una cancelación por medio de una frase de cancelación, el sistema puede borrar los valores de franjas  
30 aceptados en el intercambio anterior, emitir un mensaje guía como, por ejemplo, "Lo siento, Inténtelo otra vez", y repetir el intercambio anterior. Alternativamente, el sistema puede tratar la respuesta del usuario pronunciada en el turno de usuario como una corrección y repetir el turno de confirmación/aclaración, por ejemplo, con mensajes guía del tipo "Perfecto. De Boston a San Francisco. ¿Es correcto?". Si el usuario pronuncia entonces una frase de aceptación tal como "Sí" en respuesta, se repite el mensaje guía de encabezamiento para el nuevo intercambio.  
35 Como es evidente, puede usarse una confirmación activa después de que el usuario emita una frase de cancelación en respuesta a una confirmación pasiva. Sin embargo, pueden usarse análogamente otros tipos de confirmación después de que el usuario emita una frase de cancelación en respuesta a una confirmación pasiva. Si el siguiente intercambio contiene un mensaje guía para una respuesta de "Sí/No", puede desactivarse la funcionalidad de volver  
40 atrás a una confirmación pasiva anterior para la confirmación pasiva.

**[0035]** En una desambiguación, el sistema emite un mensaje guía que contiene una lista de las correspondencias principales y solicita al usuario que seleccione una de las correspondencias principales, por ejemplo, por su número correspondiente. Cada correspondencia puede incluir un único valor de franja o un conjunto de valores de franjas y puede ser presentada por un mensaje guía similar a un mensaje guía de confirmación que  
45 contiene todos los valores de franjas reconocidos. Cuando se selecciona una correspondencia, el valor o valores correspondientes se tratan como si hubieran sido pronunciados en el turno de usuario primario y el sistema repite el proceso de confirmación/aclaración.

**[0036]** El sistema puede también notificar al usuario un no reconocimiento. En particular, cuando un turno de  
50 usuario no es reconocido por el sistema, el sistema puede emitir un mensaje guía de excepción tal como "No lo he entendido" y repetir el turno anterior del sistema o una versión corregida del mismo.

**[0037]** El sistema puede configurarse con diversas características y funcionalidades para facilitar adicionalmente el intercambio tal como se describirá después en más detalle con referencia a las **FIG. 11-13**. Por  
55 ejemplo, el sistema puede configurarse para mantener una lista con saltos en un intercambio dado en el que se añade una entrada de la lista con saltos correspondiente al conjunto de valores de franjas presentado en una confirmación cada vez que un usuario cancela la confirmación durante ese intercambio. La lista con saltos ayuda a garantizar que, dentro del mismo intercambio, el sistema no usa y así no presentará de nuevo un conjunto de valores que coincide con una entrada de la lista con saltos. En su lugar, el sistema puede usar la siguiente mejor

correspondencia cuando esté disponible.

**[0038]** La **FIG. 11** es un organigrama que ilustra un ejemplo de una conversación que incluye otra funcionalidad opcional del sistema, que es una funcionalidad de retroceso. Específicamente, el usuario puede pronunciar una orden para volver atrás, por ejemplo, "volver atrás", en cualquier momento de manera que se regrese al inicio del turno anterior, al inicio del intercambio actual o al inicio del segmento actual, dependiendo de la aplicación. Si el sistema vuelve a una etapa que tiene algunas franjas llenas, estas franjas pueden borrarse.

**[0039]** El sistema de diálogo multifranja puede configurarse para presentar de forma adaptativa un mensaje guía más directo después de que haya tenido lugar un número umbral de excepciones, correcciones o retrocesos durante un intercambio. Por ejemplo, el sistema puede presentar un mensaje guía de encabezamiento nuevo y más directo "¿En qué fecha recogerá usted el vehículo? Por ejemplo, diga mañana, el próximo martes o el 24 de julio". Si siguen produciéndose excepciones, correcciones o retrocesos; el sistema puede transferir al usuario a un operador humano para continuar con la asistencia.

**[0040]** La **FIG. 12** es un organigrama que ilustra un ejemplo de una conversación que incluye otra funcionalidad opcional del sistema, que es una funcionalidad de cambio. La funcionalidad de cambio permite al usuario cambiar un valor de franja llena anteriormente pronunciando, por ejemplo, "Quiero cambiar el aeropuerto". Si la declaración del usuario que solicita un cambio no llena todas las franjas necesarias para una solicitud de cambio, el sistema inicia un intercambio de seguimiento como, por ejemplo, "¿El aeropuerto de salida o el de destino?". La orden de cambio puede ser confirmada opcionalmente usando cualquiera de los mecanismos de confirmación descritos anteriormente. Por ejemplo, el sistema puede confirmar activamente la orden de cambio con mensajes guía como "He entendido que desea usted modificar el aeropuerto de llegada. ¿Es correcto?"

**[0041]** La orden de cambio cancela el intercambio actual del usuario y borra al menos algunos de los valores de franjas rellenos anteriormente. Se inicia entonces un nuevo intercambio que solicita al usuario que rellene las franjas borradas en uno o más intercambios. Una vez llenas las franjas borradas, el sistema prosigue con el procesamiento y sorteará algunos intercambios anteriores si las franjas asociadas con esos intercambios siguen todavía llenas.

**[0042]** La **FIG. 13** es un organigrama que ilustra un ejemplo de una conversación que incluye otra funcionalidad opcional del sistema, es decir, una funcionalidad de revisión. En particular, el usuario puede solicitar una revisión de un valor de franja llenado anteriormente pronunciando una frase de solicitud tal como "¿Puedo comprobar mi fecha de salida?". En respuesta a la orden de revisión, el sistema reproduce el valor de franja llenado, como, por ejemplo, "Va usted a volar el viernes, 9 de agosto de 2002" y regresa al intercambio actual. Si el usuario no proporciona los valores para todas las franjas necesarias para una solicitud de revisión, el sistema inicia un intercambio de seguimiento tal como "¿Fecha de salida o de regreso?"

#### Implementación de un diálogo multifranja

**[0043]** A continuación se describirá un sistema o plataforma de ejemplo para implementar un diálogo multifranja. Simplemente a modo de ejemplo, la plataforma para implementar diálogos multifranja se describe en la presente memoria descriptiva como una implementación con uso de Java. Sin embargo, debe entenderse que el sistema puede implementarse usando cualquier lenguaje de programación adecuado, preferentemente un lenguaje de programación orientado a objetos como Java o C++. El sistema incluye generalmente objetos de franja, objetos de grupos de franjas, objetos de segmento y objetos de diálogo. A continuación se describe cada uno de estos objetos.

**[0044]** Un objeto de franja representa una franja que, tal como se observa anteriormente, es un elemento de información que se obtendrá del usuario. Un objeto de franja contiene la gramática que reconoce los posibles valores que pueden ser proporcionados por el usuario para la franja correspondiente, las reglas que buscan correspondencias para los resultados de reconocimiento de gramática para valores semánticos para la franja y las variables que indican el estado activado y lleno (entre otras variables de estado) para la franja. Los objetos de franja pueden basarse en una interfaz Java que proporciona la funcionalidad básica por omisión y/o la funcionalidad común a todos los objetos de franja.

**[0045]** La gramática que reconoce los posibles valores que pueden ser proporcionados por el usuario para la franja correspondiente es una especificación formal de las declaraciones que el sistema acepta para expresar los posibles valores para la franja. La gramática puede incluir el vocabulario de palabras que pueden usarse y

estructuras válidas para secuenciar esas palabras. Por ejemplo, la gramática para una franja de fecha debería permitir el reconocimiento de varios formatos de fecha, comprendidos entre especificaciones absolutas tales como "el doce de enero de dos mil cuatro" a formas relativas tales como "el próximo viernes" y términos familiares como "hoy" y "ayer". La gramática puede incluir también rellenos que pueden preceder y/o ser posteriores al valor de franja en una declaración de usuario pero que no especifican o distinguen un valor de franja de otro. Por ejemplo, una franja de aeropuerto puede tener la frase de relleno anterior "Volaré desde". Algunas gramáticas pueden muy específicas de la aplicación como, por ejemplo, la gramática para los aeropuertos admitidos en una aplicación de reservas de vuelos. Otras gramáticas pueden reutilizarse en varias aplicaciones, como, por ejemplo, la gramática para una fecha, una hora o un importe económico. La parte común de estas gramáticas puede predefinirse en un objeto de gramática y adaptarse a medida para una franja en particular.

**[0046]** Además de las reglas de la gramática, cada objeto de franja también contiene reglas que buscan correspondencias para los resultados de reconocimiento de gramática con los valores semánticos para la franja correspondiente que tienen significado para la aplicación específica. Por ejemplo, un objeto de franja del aeropuerto de destino puede buscar correspondencias para las declaraciones "San Francisco", "Aeropuerto San Francisco", "Aeropuerto Internacional de San Francisco" y "SFO" en un mismo identificador de aeropuerto como, por ejemplo, "SFO". Como otro ejemplo, un objeto de franja de fecha puede buscar correspondencias para la declaración "mañana" con un valor de fecha que se calcula como la fecha posterior a la fecha actual.

**[0047]** Cada objeto de franja mantiene también una serie de variables de estado o señalizaciones usadas durante la ejecución de un diálogo multifranja para indicar el estado activado y lleno (entre otras variables de estado) para la franja correspondiente. Entre los ejemplos de señalizaciones se incluyen señalizaciones activada, opcional, llena y pendiente. En particular, una señalización activada se fija como valor verdadero para indicar que una franja puede ser aceptada en el turno de usuario inminente. Una señalización opcional se fija como valor verdadero si no es necesario que el usuario proporcione explícitamente una franja activada. Una señalización llena se fija como valor verdadero cuando se ha aceptado un valor de franja después de cualquier confirmación/aclaración aplicable. Una señalización pendiente se fija como valor verdadero si un valor para la franja ha sido reconocido por el usuario pero todavía no ha sido aceptado como, por ejemplo, en una confirmación pendiente.

**[0048]** El sistema mantiene objetos de grupos de franjas para cada segmento de franjas, algunos de cuyos ejemplos se muestran en las **FIG. 14** y **15** para un segmento de itinerario de vuelo. Cada objeto de grupos de franjas define parámetros o propiedades asociados con un grupo en particular o una combinación de franjas en el segmento de franjas. Entre los ejemplos de propiedades de grupos de franjas se incluyen mensajes guía tales como mensajes guía principales o de encabezamiento, mensajes guía de ayuda, mensajes guía de error, mensajes guía de confirmación y/o mensajes guía de desambiguación, así como propiedades de reconocimiento, es decir, variables que influyen en la conducta de reconocimiento tales como tiempos de desconexión, umbrales de reconocimiento, parámetros de reconocimiento, políticas de memorias intermedias rápidas y así sucesivamente. Pueden definirse diferentes clases de grupos de franjas como, por ejemplo, clases de grupos de franjas de reconocimiento anterior y reconocimiento posterior tal como se muestra en las **FIG. 14** y **15**, respectivamente, para mantener conjuntos de propiedades separados.

**[0049]** Específicamente, la **FIG. 14** ilustra algunos de los posibles objetos de grupos de franjas para la clase de grupos de franjas de reconocimiento anterior para el segmento de itinerario de vuelo. La clase de grupos de franjas de reconocimiento anterior puede contener los mensajes guía y las propiedades de reconocimiento usados antes de que se reconozcan los valores de franjas, tales como mensajes guía principales o de encabezamiento, mensajes guía de ayuda, mensajes guía de no correspondencia, mensajes guía de entrada, tiempos de desconexión, umbrales de confianza y así sucesivamente. Cuando se usa la clase de grupos de franjas de reconocimiento anterior, la combinación de franjas se compararía normalmente con el conjunto de franjas activadas actualmente.

**[0050]** La **FIG. 15** ilustra algunos de los posibles objetos de grupos de franjas para la clase de grupos de franjas de reconocimiento posterior para el segmento de itinerario de vuelo. La clase de grupos de franjas de reconocimiento posterior puede contener los mensajes guía y propiedades de reconocimiento usados después de que se reconozcan los valores de franjas, tales como mensajes guía de confirmación activa, mensajes guía de desambiguación, mensajes guía de confirmación pasiva, y así sucesivamente. Cuando se usa la clase de grupos de franjas de reconocimiento posterior, la combinación de franjas se compararía normalmente con el conjunto de franjas actualmente pendientes, es decir, las franjas reconocidas del usuario pero todavía no confirmadas.

**[0051]** Cuando se requiere un parámetro como, por ejemplo, un mensaje guía o una propiedad de

reconocimiento en un momento dado de un diálogo multifranja, el sistema identifica un objeto de grupo de franjas de mejor correspondencia a partir de una clase de grupos de franjas que contiene el parámetro y busca el valor del parámetro deseado en el objeto de grupos de franjas de mejor correspondencia identificado. Por ejemplo, cuando se solicita un mensaje guía de encabezamiento en un momento dado de un diálogo multifranja con el fin de que el sistema produzca a partir del usuario los valores para las franjas en la combinación de franjas activadas actualmente, por ejemplo, fecha, hora y mañana-tarde, el sistema identifica un objeto de grupos de franjas de mejor correspondencia a partir de la clase de grupos de franjas de reconocimiento anterior en la **FIG. 14**. El sistema selecciona el objeto de grupos de franjas cuya combinación de franjas es la más próxima a la combinación de franjas activadas actualmente. La combinación de franjas más cercana puede determinarse usando varios procedimientos adecuados tales como el máximo número de franjas solapadas, el menor número de franjas no solapadas en el objeto de grupos de franjas o el menor número de franjas no solapadas en el grupo activado. En el ejemplo actual, el sistema identifica y usa el objeto de grupos de franjas en la clase de grupos de franjas de reconocimiento anterior mostrada en la **FIG. 14** que tiene una combinación de franjas de correspondencia exacta con la combinación de franjas activadas actualmente, es decir, fecha, hora y mañana-tarde. Sin embargo, si no se encuentra una correspondencia exacta de objeto de grupos de franjas, puede identificarse y usarse otro objeto de grupo que se considere el más cercano, por ejemplo, un objeto de grupo con una combinación de dos franjas tales como las franjas de hora y de mañana-tarde. El sistema puede definir un objeto de grupos de franjas separado para cada franja del segmento de franjas con el fin de garantizar que siempre pueda encontrarse un grupo de franjas para cualquier combinación de franjas activadas.

**[0052]** El sistema incluye también objetos de segmento. Un objeto de segmento mantiene un conjunto de franjas en un segmento de franjas que determina una serie de uno o más intercambios multifranja sucesivos. Los valores para las franjas en un segmento de franjas dado puede obtenerse potencialmente del usuario en un único intercambio. Alternativamente, si el usuario no proporciona los valores para todas las franjas en ese segmento en un intercambio, se inician intercambios de seguimiento hasta que todas las franjas requeridas, es decir, no opcionales, están llenas. Cuando un segmento se llena completamente, se invoca al siguiente segmento, si existe.

**[0053]** El sistema incluye además objetos de diálogo que definen el flujo de diálogo. Aunque cada diálogo multifranja puede realizar una función diferente, el flujo de diálogo para cada diálogo tiene generalmente una estructura común. La **FIG. 16** muestra un diagrama de bloques que ilustra una estructura de flujo de diálogo de ejemplo. Después de que el sistema inicia un diálogo en el bloque 1601, el sistema obtiene el primer segmento en el diálogo en el bloque 1602. El sistema determina las franjas en este segmento que deberían ser activadas en el bloque 1603 como, por ejemplo, incluyendo las franjas que todavía no están llenas. Se selecciona el grupo de franjas de reconocimiento anterior más cercano para recuperar el mensaje guía de encabezamiento y otros parámetros de reconocimiento anterior en el bloque 1604 (como el mensaje guía de ayuda y cualquier mensaje guía de excepción). El mensaje guía de encabezamiento se pronuncia habitualmente de manera que el usuario puede proporcionar potencialmente una serie de franjas deseadas en una declaración. Por ejemplo, si las franjas incluyen una franja de fecha, una franja de hora y una franja de meridiano (mañana-tarde), el mensaje guía podría ser "¿Cuándo recogerá usted el vehículo?". A continuación, la respuesta del usuario es recibida y procesada por el sistema en el bloque 1605, por ejemplo, obteniendo las mejores hipótesis de la secuencia de palabras prevista por el usuario a partir del motor de reconocimiento del habla y realizando todas las confirmaciones, aclaraciones o desambiguaciones deseadas basándose en los ajustes personalizados de la aplicación hasta que el sistema acepte un único conjunto de valores de franjas.

**[0054]** El sistema determina a continuación si el segmento de franjas contiene o no más franjas que deban llenarse en el bloque de decisión 1606 que normalmente incluye franjas cuyos valores todavía no se han llenado. Debe observarse que el sistema, al determinar si el segmento de franjas contiene o no más franjas que deban llenarse en el bloque de decisión 1606, puede aplicar una lógica específica de la aplicación que especifique que es posible llenar automáticamente algunos valores de franjas a partir de valores de franjas, que algunas franjas son opcionales o que es preciso llenar algunas franjas adicionales como consecuencia de los valores de otras franjas aceptadas. Si quedan algunas franjas por llenar según se determina en el bloque de decisión 1606, el sistema regresa al bloque 1603 para activar el siguiente conjunto de franjas. En caso contrario, el sistema continúa al bloque de decisión 1607 para determinar si existen más segmentos de franjas. Si existe al menos un segmento de franjas más, el sistema obtiene el siguiente segmento en el bloque 1608 y regresa al bloque 1603. Alternativamente, si no existen más segmentos de franjas, el diálogo termina en el bloque 1610.

**[0055]** Cuando se procesa la entrada del usuario, el sistema puede estar en un modo normal o un modo de confirmación pasiva. El modo de confirmación pasiva se activa cuando el sistema determina que la respuesta del usuario debe confirmarse pasivamente basándose en algunos criterios predefinidos. Un conjunto de criterios de este

tipo podría ser que el nivel de confianza devuelto por el motor de habla estuviera por debajo del umbral para una aceptación en silencio pero por encima del umbral que requiere confirmación activa. Si no se activa el modo de confirmación pasiva, el sistema puede estar en modo normal por omisión. La FIG. 17 es un organigrama que ilustra un procesamiento de ejemplo de la entrada del usuario en modo normal mientras que la FIG. 18 es un organigrama que ilustra un procesamiento de ejemplo de la entrada del usuario en modo de confirmación pasiva. La determinación como modo normal o pasivo puede basarse, por ejemplo, en un nivel de confianza promediado o ponderado para el conjunto de franjas a la que corresponde la respuesta del usuario. Cuando se trata una respuesta del usuario con valores para múltiples franjas, la determinación puede realizarse con una única determinación para todas las franjas en la respuesta del usuario o como una determinación separada para cada franja en la respuesta del usuario y, por ejemplo, los resultados separados pueden promediarse o ponderarse.

**[0056]** En el modo normal tal como se ilustra en la **FIG. 17**, el motor de habla reconoce la entrada del usuario y devuelve una o más hipótesis al bloque 1701. Si el motor de habla devuelve varias posibles correspondencias tal como se determina en el bloque de decisión 1702, las posibles correspondencias pueden desambiguarse en el bloque 1708. Por ejemplo, puede presentarse al usuario una lista de correspondencias principales y pedirle que seleccione una. Cada correspondencia puede presentarse mediante un mensaje guía de desambiguación similar a un mensaje guía de confirmación. Cuando se recibe una correspondencia seleccionada por el usuario y se procesa en el bloque 1709, el sistema avanza al siguiente intercambio en el bloque 1710.

**[0057]** Alternativamente, si el motor de habla devuelve sólo una hipótesis, es decir, la mejor correspondencia, tal como se determina en el bloque de decisión 1702, el sistema determina si el nivel de confianza para la hipótesis está o no en o por encima de un umbral de confianza alto predeterminado en el bloque de decisión 1703. Si el nivel de confianza alto está en o por encima del umbral de confianza alto, el sistema acepta los valores de franjas y entra en el modo de confirmación pasiva en el bloque 1707. Alternativamente, si el nivel de confianza alto está por debajo del umbral de confianza alto, el sistema confirma activamente la mejor correspondencia emitiendo un mensaje guía de confirmación en el bloque 1704. Por ejemplo, el sistema puede construir el mensaje guía de confirmación identificando el grupo de franjas en la clase de grupos de franjas de reconocimiento posterior que sea más próximo al grupo de franjas en la mejor correspondencia y recuperando el mensaje guía correspondiente de confirmación activa. La respuesta del usuario al mensaje guía de confirmación es recibido y procesado en el bloque 1705. Si el usuario cancela la confirmación, por ejemplo, pronunciando una frase de cancelación tal como "no" o "no es correcto", y proporciona una corrección, los valores corregidos pueden regresar al bloque 1704 para su confirmación. Si el usuario cancela la confirmación sin corrección, se repite el intercambio actual en el bloque 1706. En otras palabras, se descartan o se borran posibles valores de franjas pendientes y el sistema repite el mensaje guía de encabezamiento que se reprodujo originalmente. Si el usuario acepta la confirmación, por ejemplo, pronunciando una frase tal como "sí" o "es correcto", el sistema acepta los valores de franjas y continúa con el siguiente intercambio en el bloque 1710.

**[0058]** En el modo de confirmación pasiva tal como se ilustra en la **FIG. 18**, el mensaje guía de confirmación de un intercambio se combina con el mensaje guía de encabezamiento para el siguiente intercambio. Por ejemplo, el mensaje guía combinado puede ser "Perfecto. Boston. ¿Y en qué fecha va usted a volar?". La respuesta del usuario es reconocida a continuación por el motor de habla en el bloque 1801. Si la respuesta del usuario no empieza con una frase de aceptación o de cancelación, es decir, un sí o un no, tal como se determina en el bloque de decisión 1802, la respuesta del usuario se procesa como en el caso del modo normal descrito anteriormente. Alternativamente, si la respuesta del usuario empieza con una frase de aceptación o de cancelación tal como se determina en el bloque de decisión 1802, pueden verse afectados los valores de franjas anteriores. Si la respuesta del usuario es una respuesta afirmativa del usuario, el mensaje guía de encabezamiento para el nuevo intercambio puede simplemente repetirse, por ejemplo "¿Y en qué fecha va usted a volar?" en el bloque 1803. Si la respuesta del usuario es una respuesta negativa del usuario sin ninguna corrección, el sistema puede borrar los valores de franjas aceptados en el intercambio anterior, reproducir un mensaje tal como "Lo siento, Inténtelo otra vez", repetir el intercambio anterior en el bloque 1804. Para una respuesta negativa del usuario con una corrección, los valores corregidos son confirmados explícitamente en el bloque 1805 hasta que se aceptan o se cancelan explícitamente las franjas de intercambio anteriores.

**[0059]** Durante un turno de usuario primario, el sistema permite al usuario proporcionar valores para una o más franjas activadas. La gramática para el turno debería así configurarse de forma flexible para reconocer varias combinaciones de franjas flexibles. Por ejemplo, la **FIG. 19** ilustra una posible gramática para un intercambio con tres franjas. Cada rama del diagrama representa una regla de gramática que podría coincidir con la declaración del usuario. Las gramáticas para las franjas 1, 2 y 3 se representan mediante números de referencia 1901, 1902, 1903, respectivamente. El operador de interrogación "?" aplicado a cada una de las gramáticas 1902, 1903

correspondientes a las franjas 2 y 3 en la primera rama indica que las franjas 2 y 3 son opcionales de manera que en esta primera rama la declaración del usuario contiene un valor para la franja 1, opcionalmente un valor para la franja 2, y, si existe un valor para la franja 2, opcionalmente un valor para la franja 3. Debe observarse que una aplicación puede restringir la combinación permitida de franjas (tal como una franja que tiene que llegar antes que otra) dependiendo de las normas del lenguaje usado y del contexto de un intercambio en particular.

5  
10  
15  
[0060] Durante un turno de confirmación, el sistema permite al usuario no sólo aceptar o cancelar la confirmación, sino también proporcionar valores corregidos. La gramática para un turno de confirmación debería construirse así para incluir las gramáticas de aceptación y cancelación así como gramáticas para las franjas que se confirman, un ejemplo de las cuales se muestra en la **FIG. 20**. La gramática de aceptación 2001 contiene un conjunto de frases que expresan una respuesta afirmativa, tales como "sí", "está bien" y "correcto". La gramática de cancelación 2002 contiene un conjunto de frases que expresan una respuesta negativa, tales como "no", "no es correcto" e "incorrecto". Si la gramática de cancelación está presente se incluye una gramática de 2003 opcional que reconoce nuevos valores para las franjas canceladas.

20  
[0061] Aunque en la presente invención se describen e ilustran realizaciones de ejemplo en la presente memoria descriptiva, se observará que son meramente ilustrativas y que es posible introducir modificaciones a estas realizaciones sin apartarse del ámbito de la invención. Por ejemplo, aunque los sistemas y procedimientos multifranja descritos en la presente memoria descriptiva están bien adaptados para interacciones de voz que usan sistemas de reconocimiento del habla, los sistemas y procedimientos multifranja pueden adaptarse también para su uso con interacciones multifranja basadas en texto tales como un diálogo multifranja interactivo basado en Internet. Así, el ámbito de la invención pretende definirse sólo en términos de las siguientes reivindicaciones.

## REIVINDICACIONES

1. Un procedimiento implementado en ordenador para construir y procesar un diálogo multifranja con un usuario, que comprende las etapas de:
  - 5 activación de todas las franjas no llenas en un segmento de franjas en el que una franja es un elemento de datos cuyo valor puede obtenerse de una entrada de usuario, y en el que una señalización activada se fija como valor verdadero para indicar que una franja puede ser aceptada en un turno de usuario inminente;
  - 10 emisión de un mensaje guía de encabezamiento primario para producir valores para una o más franjas activadas; recepción de una cualquiera de una respuesta de usuario primaria basada en texto o basada en el habla en el mensaje guía de encabezamiento primario que produce valores para las franjas activadas;
  - 15 procesamiento de la respuesta de usuario primaria para determinar al menos un posible valor de reconocimiento para cada franja contenida en la respuesta de usuario primaria; llenado de cada franja contenida en la respuesta de usuario primaria con un valor de correspondencia seleccionado entre al menos un posible valor de reconocimiento correspondiente; y
  - 20 repetición de la activación, la emisión, la recepción, el procesamiento y el llenado para todas las franjas no llenas en el segmento hasta que se llenen todas las franjas en el segmento de franjas.
2. El procedimiento según la reivindicación 1, que comprende además:
  - 25 la realización de turnos en al menos uno entre confirmar y aclarar los valores de franjas de correspondencia para las franjas contenidas en la respuesta de usuario primaria.
  3. El procedimiento según la reivindicación 2, en el que el al menos uno entre confirmar y aclarar se
  - 30 selecciona entre el grupo que consiste en aceptar en silencio una mejor correspondencia, confirmar pasivamente la mejor correspondencia, confirmar activamente la mejor correspondencia, desambiguar entre las mejores correspondencias y notificar al usuario un no reconocimiento.
  4. El procedimiento según la reivindicación 3, en el que el al menos uno entre confirmar y aclarar se
  - 35 selecciona basándose en el número de posibles valores de reconocimiento para las franjas en la respuesta de usuario primaria y un nivel de confianza correspondiente para cada uno de los posibles valores de reconocimiento.
  5. El procedimiento según la reivindicación 3, en el que cuando el al menos uno entre confirmar y aclarar es una confirmación activa, la realización de turnos incluye el reconocimiento de una respuesta de confirmación del
  - 40 usuario como una entre una confirmación, una cancelación y una cancelación y corrección, y en el que cuando la respuesta de confirmación del usuario es una cancelación y corrección, la corrección se procesa procesando la corrección para determinar al menos un posible valor de reconocimiento para cada franja contenida en la corrección.
  6. El procedimiento según la reivindicación 3, en el que cuando el al menos uno entre confirmar y aclarar es una confirmación pasiva, la realización de turnos incluye el reconocimiento de una respuesta pasiva de confirmación del usuario como una entre una respuesta a un mensaje guía de encabezamiento primario siguiente, una confirmación, una cancelación y una cancelación y corrección, y en el que cuando la respuesta pasiva de confirmación del usuario es una cancelación y corrección, la corrección se procesa procesando la corrección para
  - 50 determinar al menos un posible valor de reconocimiento para cada franja contenida en la corrección.
  7. El procedimiento según la reivindicación 1, en el que el procesamiento de la respuesta de usuario primaria incluye la aplicación de reglas de gramática para facilitar el reconocimiento de posibles valores para una franja correspondiente.
  - 55 8. El procedimiento según la reivindicación 1, en el que la emisión del mensaje guía de encabezamiento primario depende del conjunto de segmentos no llenados.
  9. El procedimiento según la reivindicación 1, en el que la emisión es una entre emisión basada en texto y emisión basada en el habla.

10. Un sistema informático para construir y procesar un diálogo multifranja con un usuario, proporcionando el sistema informático un sistema de aplicación de voz que comprende:
- 5 una pluralidad de objetos de franja cada uno de los cuales representa una franja en un segmento en el que una franja es un elemento de datos cuyo valor puede obtenerse de una entrada de usuario, siendo posible asignar a cada franja un valor basándose en el diálogo multifranja;
- al menos un objeto de segmento de franjas, conteniendo cada objeto de segmento de franjas un conjunto correspondiente de objetos de franja;
- 10 al menos un conjunto de objetos de grupos de franjas para cada objeto de segmento de franjas, definiendo cada objeto de grupos de franjas parámetros asociados con una combinación de franjas particular en el objeto de segmento de franjas; y
- 15 objetos de diálogo que definen un flujo del diálogo multifranja para hacer que el sistema:
- i) obtenga un primer segmento en el diálogo;
- 20 ii) determine las franjas en este segmento que deberían activarse incluyendo las franjas que todavía no están llenas, en el que una señalización activada se fija como valor verdadero para indicar que la franja puede ser aceptada en un turno de usuario inminente;
- iii) active las franjas no llenas determinadas;
- 25 iv) emita un mensaje guía de encabezamiento primario para solicitar que el usuario proporcione valores para una o más franjas activadas;
- v) reciba una respuesta de usuario primaria al mensaje guía de encabezamiento primario solicitando al usuario que proporcione valores para la una o más franjas activadas;
- 30 vi) procese la respuesta de usuario primaria obteniendo las mejores hipótesis de la secuencia de palabras prevista por el usuario a partir de un motor de reconocimiento del habla para obtener un solo conjunto de valores de franjas;
- 35 vii) determine si el segmento de franjas contiene o no más franjas que deban llenarse; y
- viii) si las franjas no están llenas, repita de ii) a vii) hasta que todas las franjas del segmento de franjas estén llenas.
11. El sistema según la reivindicación 10, en el que el sistema se implementa en un lenguaje de programación orientado a objetos.
- 40
12. El sistema según la reivindicación 10, en el que cada parámetro para al menos un conjunto de objetos de grupos de franjas se clasifica en una de al menos dos clases de grupos de franjas definidas para facilitar el mantenimiento de conjuntos separados de los parámetros.
- 45
13. El sistema según la reivindicación 12, en el que las clases de grupos de franjas incluyen una clase de grupos de franjas de reconocimiento anterior y una clase de grupos de franjas de reconocimiento posterior.
14. El sistema según la reivindicación 10, en el que cada objeto de franja contiene una gramática para facilitar el reconocimiento de un posible valor proporcionado por el usuario para la franja correspondiente, reglas que facilitan la búsqueda de correspondencias para los resultados de reconocimiento de gramática de valores semánticos para la franja correspondiente, y variables que indican un estado para la franja correspondiente.
- 50
15. El sistema según la reivindicación 10, en el que los parámetros definidos por cada objeto de grupos de franjas se seleccionan entre el grupo que consiste en mensajes guía de encabezamiento, mensajes guía de ayuda, mensajes guía de error, mensajes guía de confirmación, mensajes guía de desambiguación y propiedades de reconocimiento.
- 55
16. El sistema según la reivindicación 10, en el que el sistema es uno entre basado en texto y basado en

el habla.

17. El procedimiento según la reivindicación 1, en el que el diálogo multifranja se define mediante objetos que comprenden:

5

un objeto de segmento de franjas que contiene un conjunto correspondiente de objetos de franja que representan cada uno una franja en el segmento;

10 al menos un conjunto de objetos de grupos de franjas para cada objeto de segmento de franjas, definiendo cada objeto de grupos de franjas parámetros asociados con una combinación particular de franjas en el objeto de segmento de franjas; y

objetos de diálogo que definen las etapas del diálogo multifranja.

15 18. El procedimiento según la reivindicación 17, en el que cada objeto de franja contiene una gramática para facilitar el reconocimiento de un posible valor proporcionado por el usuario para la franja correspondiente, reglas que buscan correspondencias para los resultados de reconocimiento de gramática para valores semánticos para la franja correspondiente, y variables que indican un estado para la franja correspondiente.

20 19 El procedimiento según la reivindicación 17, en el que cada parámetro para al menos un conjunto de objetos de grupos de franjas se clasifica en una de al menos dos clases de grupos de franjas definidas para facilitar el mantenimiento de conjuntos separados de los parámetros.

20. El procedimiento según la reivindicación 19, en el que las clases de grupos de franjas incluyen una 25 clase de grupos de franjas de reconocimiento anterior y una clase de grupos de franjas de reconocimiento posterior.

21. El procedimiento según la reivindicación 17, en el que los parámetros definidos por cada objeto de grupos de franjas se selecciona entre el grupo que consiste en mensajes guía de encabezamiento, mensajes guía de ayuda, mensajes guía de error, mensajes guía de confirmación, mensajes guía de desambiguación y propiedades 30 de reconocimiento.

22. El procedimiento según la reivindicación 17, en el que el sistema es uno entre basado en texto y basado en el habla.

35 23. El procedimiento según la reivindicación 17, en el que la ejecución incluye la realización de una acción en respuesta a una entrada de usuario, seleccionándose la acción entre el grupo que consiste en aceptar en silencio una mejor correspondencia, confirmar pasivamente la mejor correspondencia, confirmar activamente la mejor correspondencia, desambiguar entre las mejores correspondencias y notificar al usuario un no reconocimiento.

40 24. El procedimiento según la reivindicación 23, en el que cuando la acción es una confirmación activa, la ejecución incluye además el reconocimiento de una respuesta de confirmación del usuario como una entre una confirmación, una cancelación y una cancelación y corrección, y cuando la respuesta de confirmación del usuario es una cancelación y corrección, la corrección se procesa procesando la corrección para determinar al menos un posible valor de reconocimiento para cada franja contenida en la corrección.

45

25. El procedimiento según la reivindicación 23, en el que cuando la acción es una confirmación pasiva, la ejecución incluye además el reconocimiento de una respuesta pasiva de confirmación del usuario como una entre una respuesta a un mensaje guía de encabezamiento primario siguiente, una confirmación, una cancelación y una cancelación y corrección, y cuando la respuesta pasiva de confirmación del usuario es una cancelación y corrección, 50 la corrección se procesa procesando la corrección para determinar al menos un posible valor de reconocimiento para cada franja contenida en la corrección.

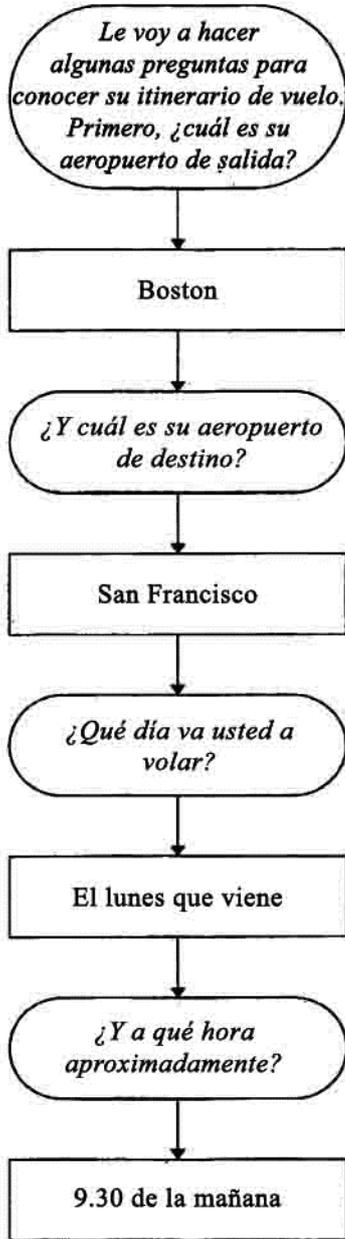
26. El procedimiento según la reivindicación 23, en el que la acción se selecciona basándose en una serie de posibles valores para las franjas y un nivel de confianza correspondiente para cada posible valor.

55

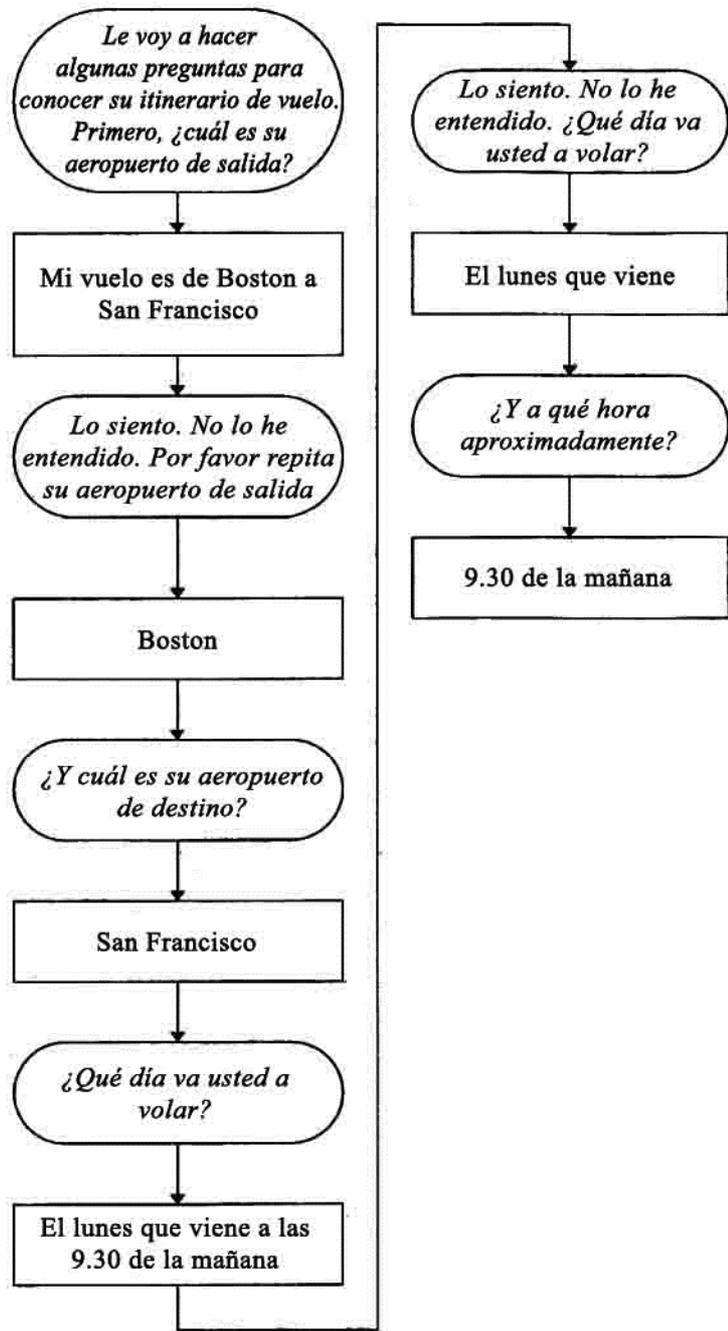
27. El procedimiento según la reivindicación 17, en el que los objetos de diálogo definen turnos entre al menos uno entre confirmar y aclarar los valores de franjas.

28. Un producto de programa informático materializado en un medio legible por ordenador, incluyendo el

producto de programa informático instrucciones que, cuando son ejecutadas por un procesador, hacen que el procesador lleve a cabo el procedimiento de cualquiera de las reivindicaciones 1 a 9 ó 17 a 27.



**TÉCNICA ANTERIOR**  
(Diálogo dirigido)  
**FIG. 1a**



**TÉCNICA ANTERIOR**  
(Diálogo dirigido)  
**FIG. 1b**

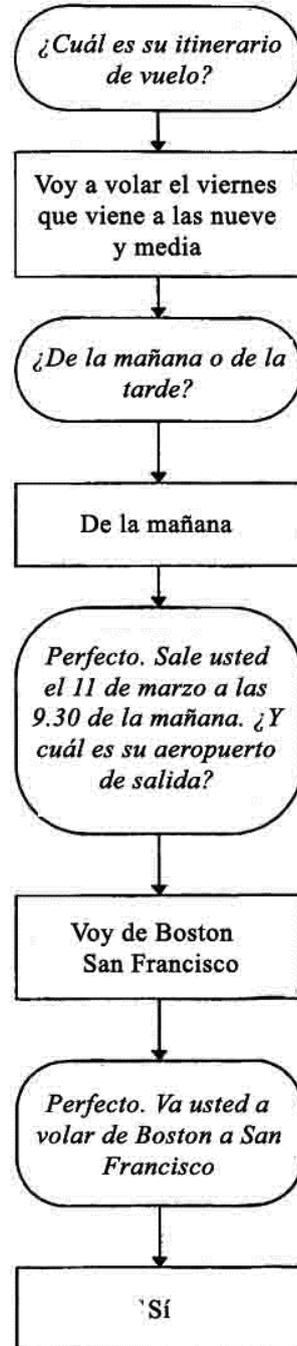
**DIÁLOGOS MULTIFRANJA**



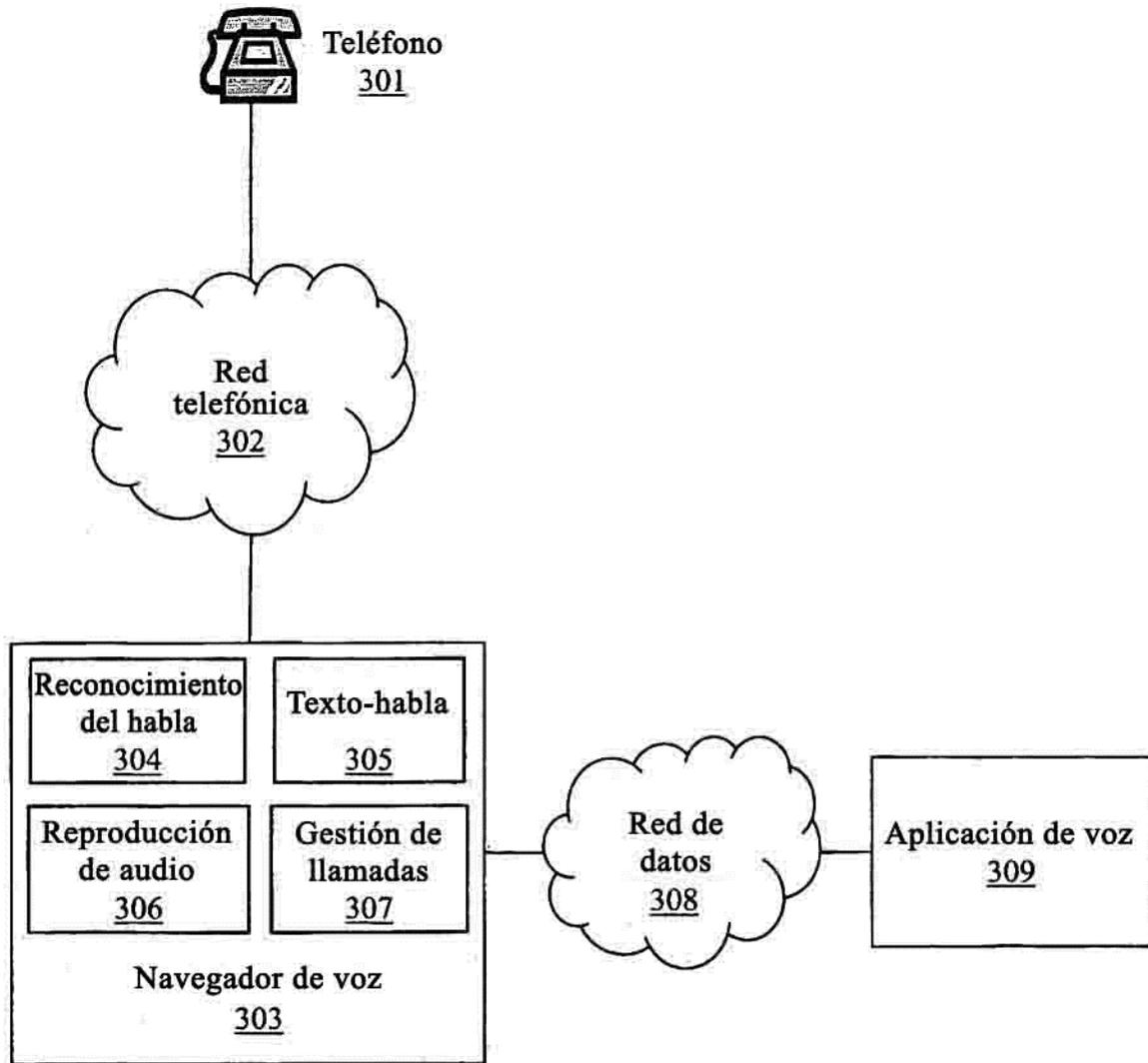
**FIG. 2a**



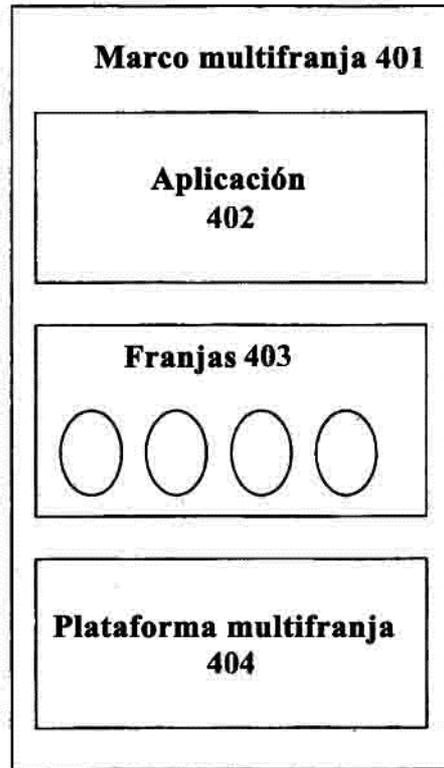
**FIG. 2b**



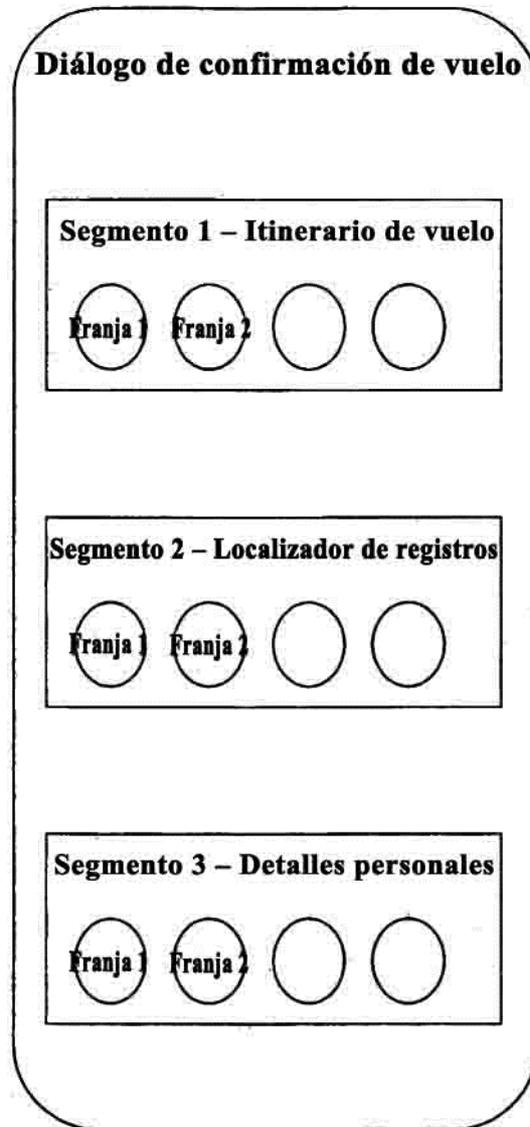
**FIG. 2c**



**ARQUITECTURA DE APLICACIÓN DE VOZ  
FIG. 3**

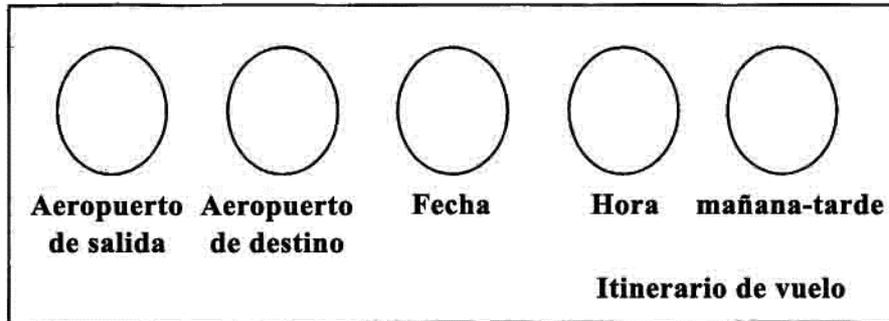


**FIG. 4**



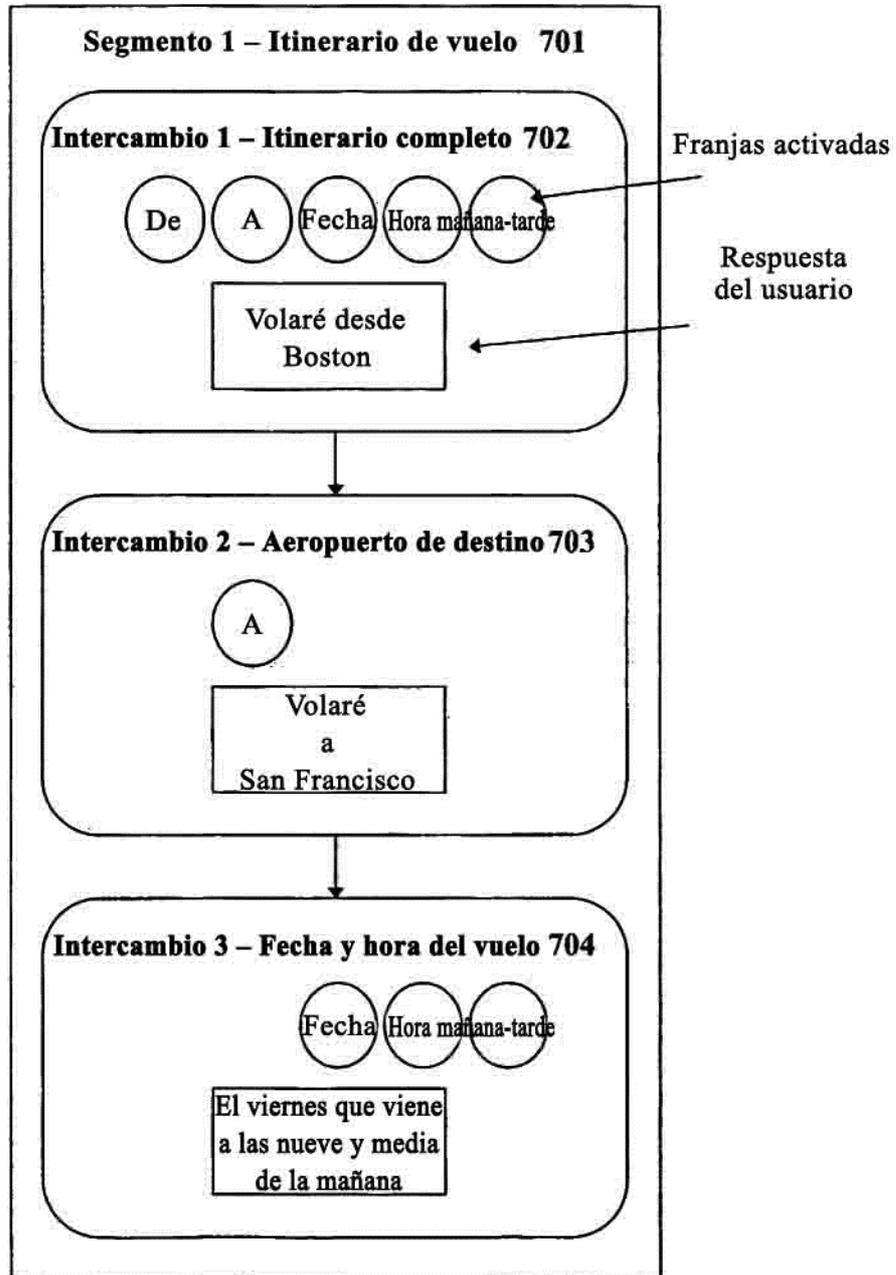
**ESTRUCTURA DE DIÁLOGO MULTIFRANJA**

**FIG. 5**



**SEGMENTO DE DIÁLOGO**

**FIG. 6**



**FLUJO DE DIÁLOGO  
FIG. 7**

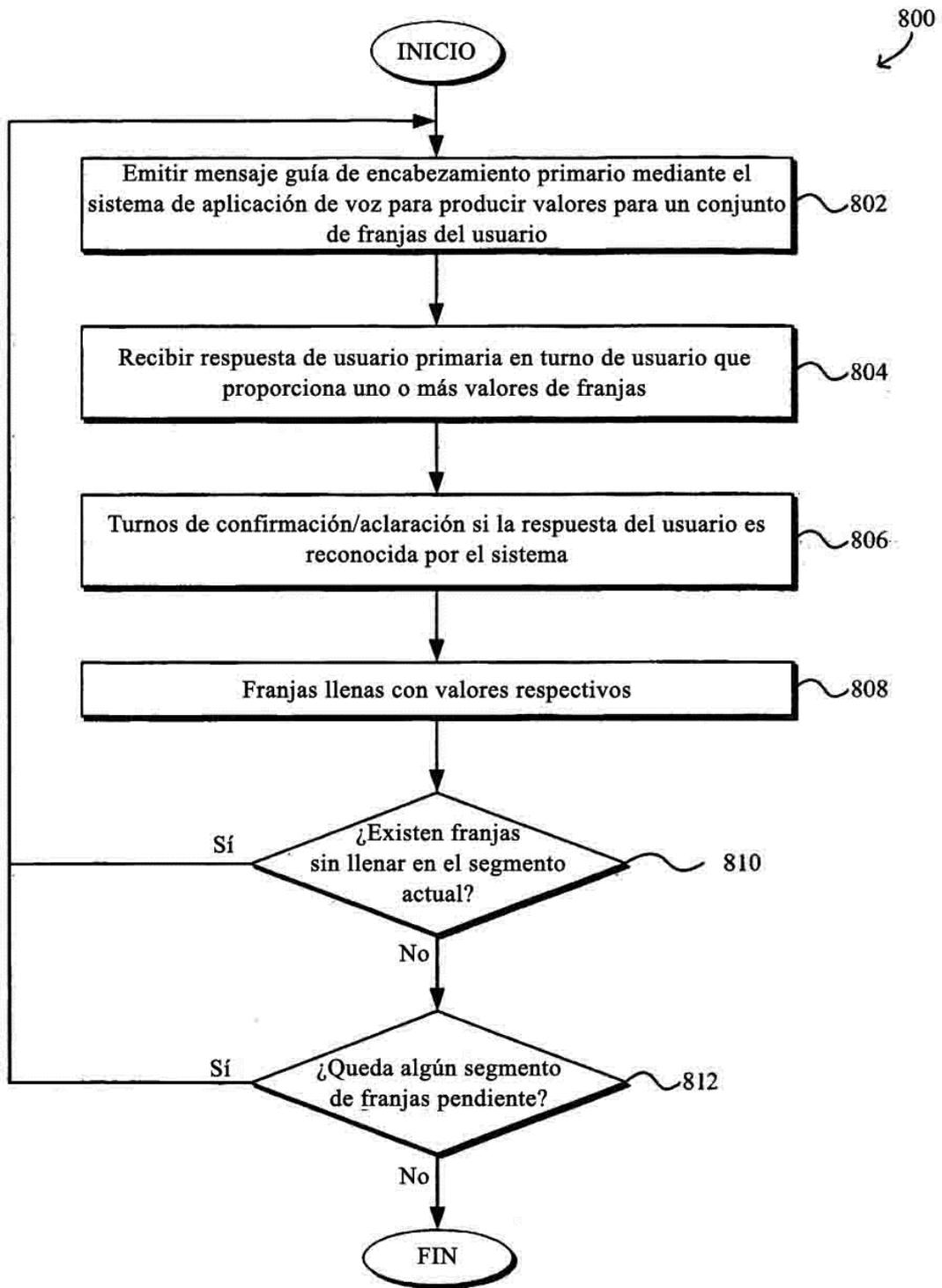
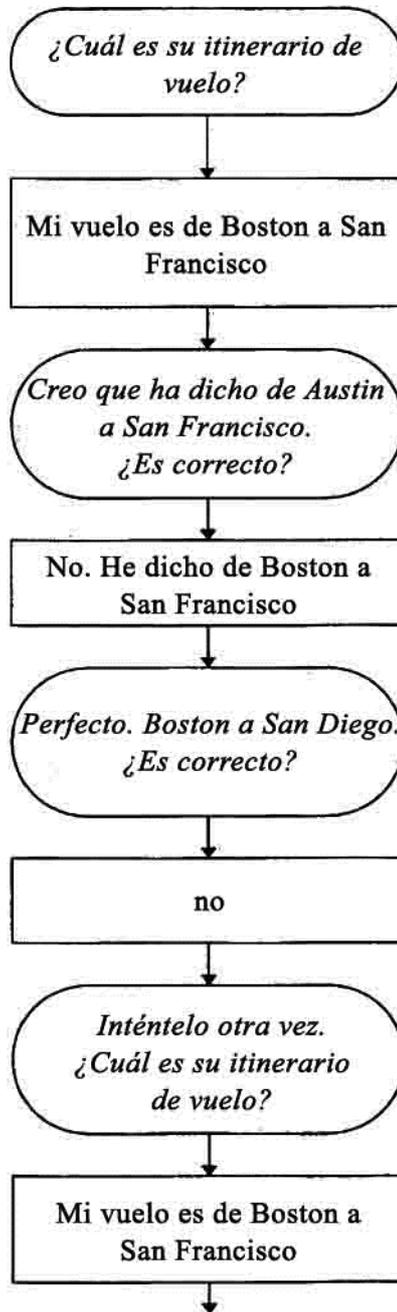
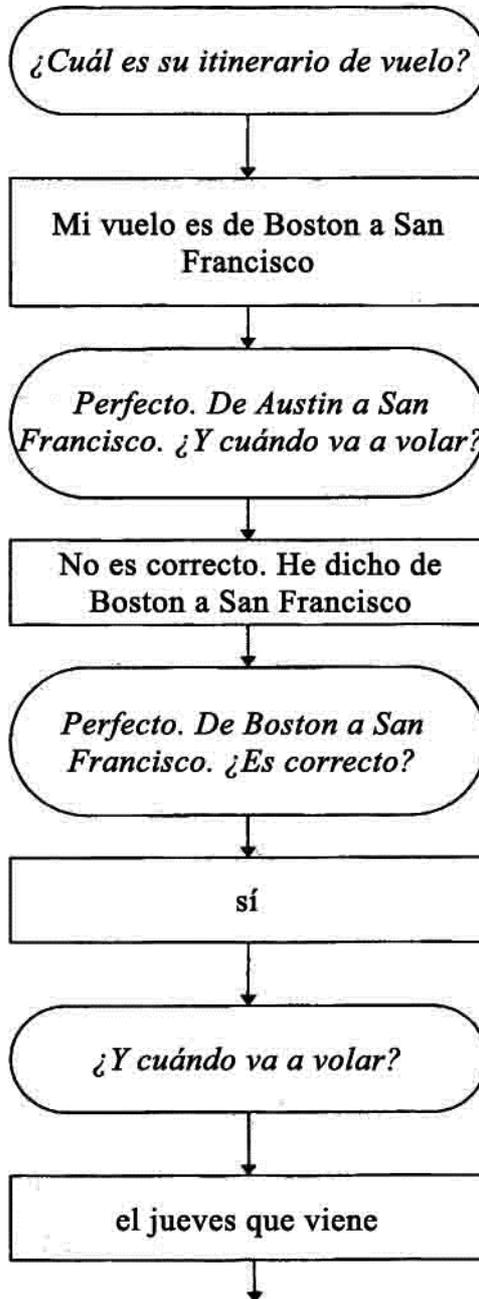


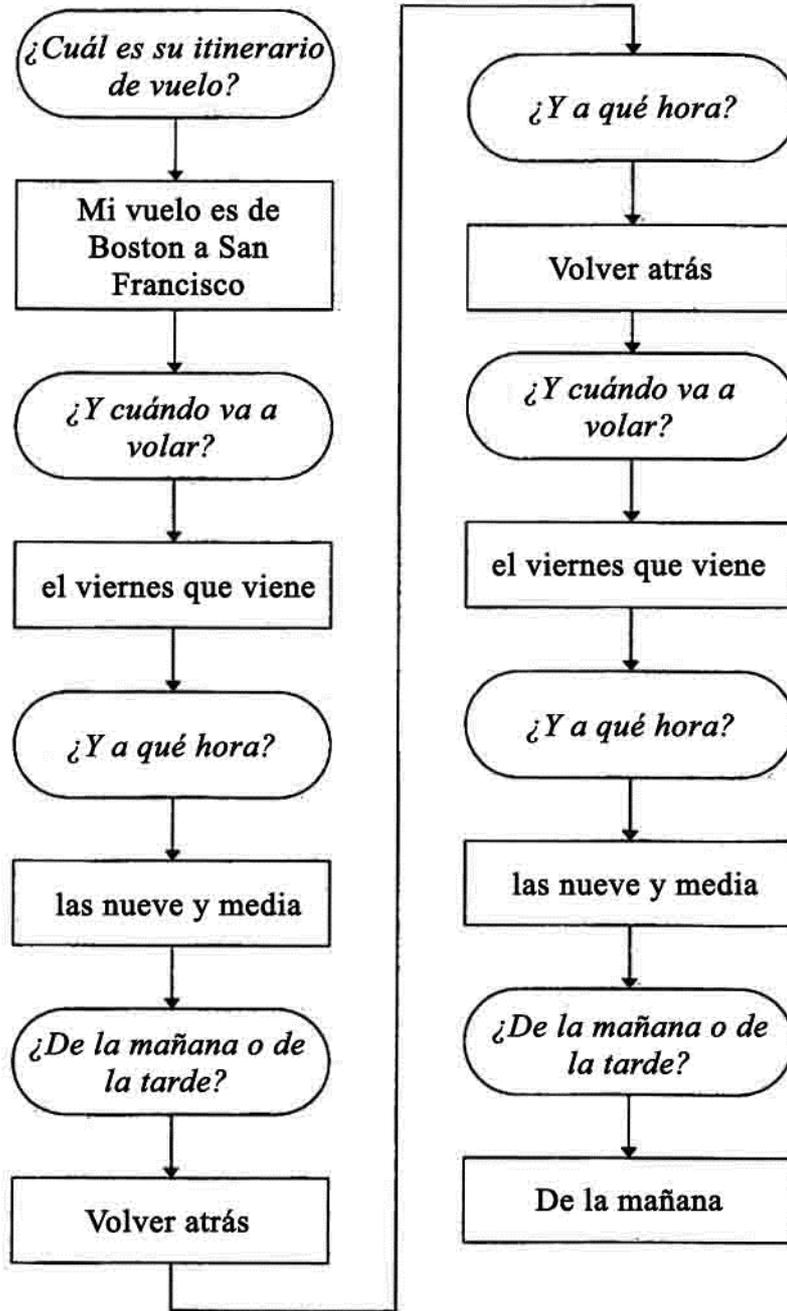
FIG. 8



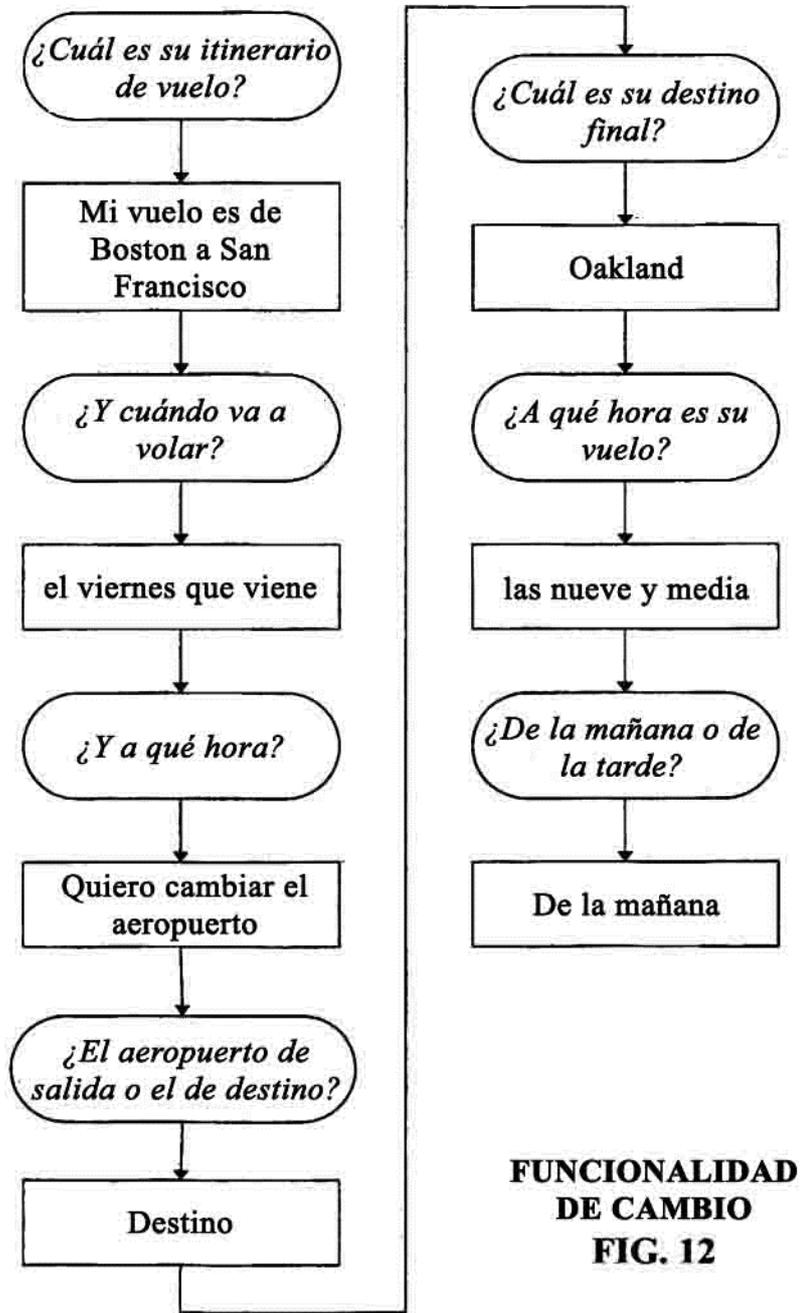
**FUNCIONALIDAD DE CONFIRMACIÓN ACTIVA**  
**FIG. 9**

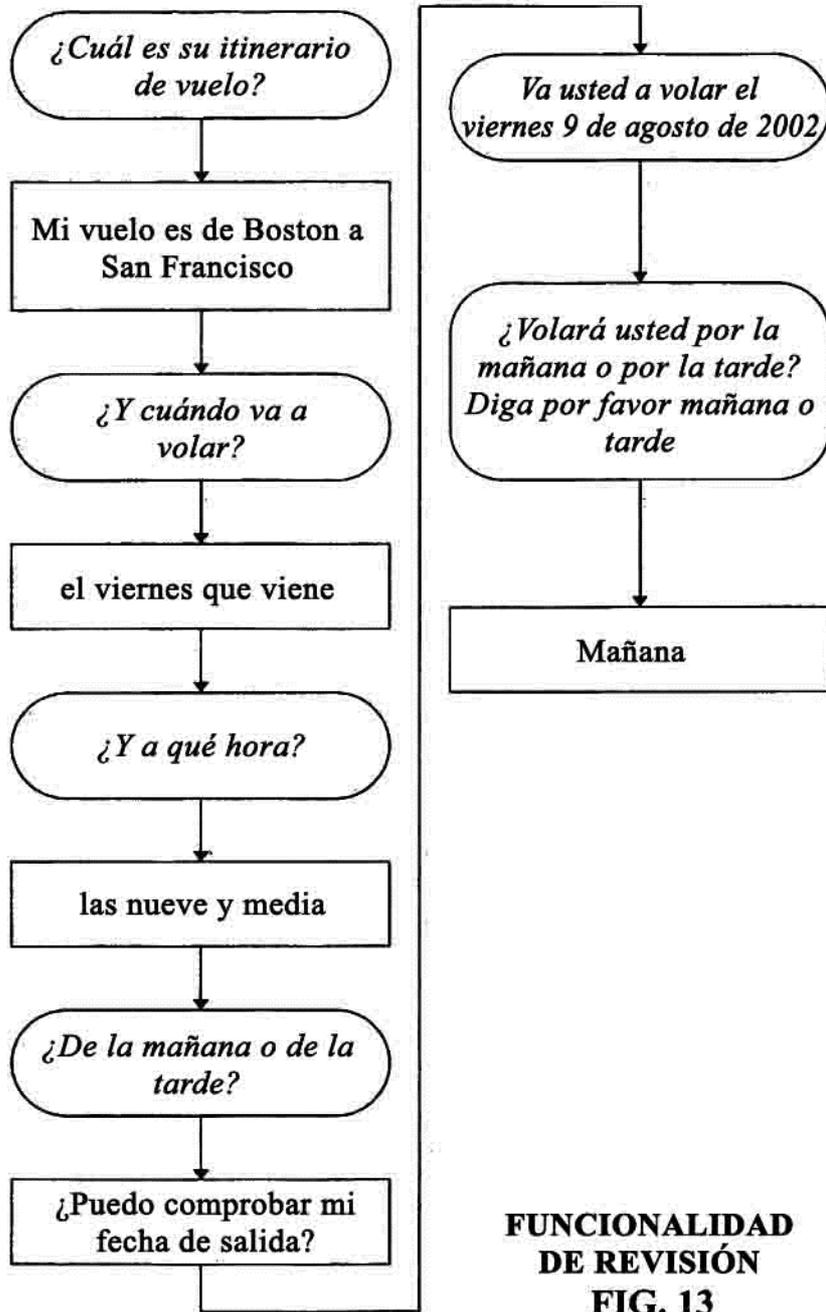


**FUNCIONALIDAD DE  
CONFIRMACIÓN PASIVA  
FIG. 10**



**FUNCIONALIDAD DE RETROCESO**  
**FIG. 11**

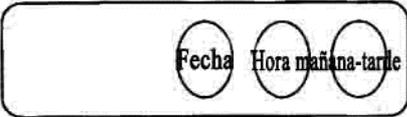




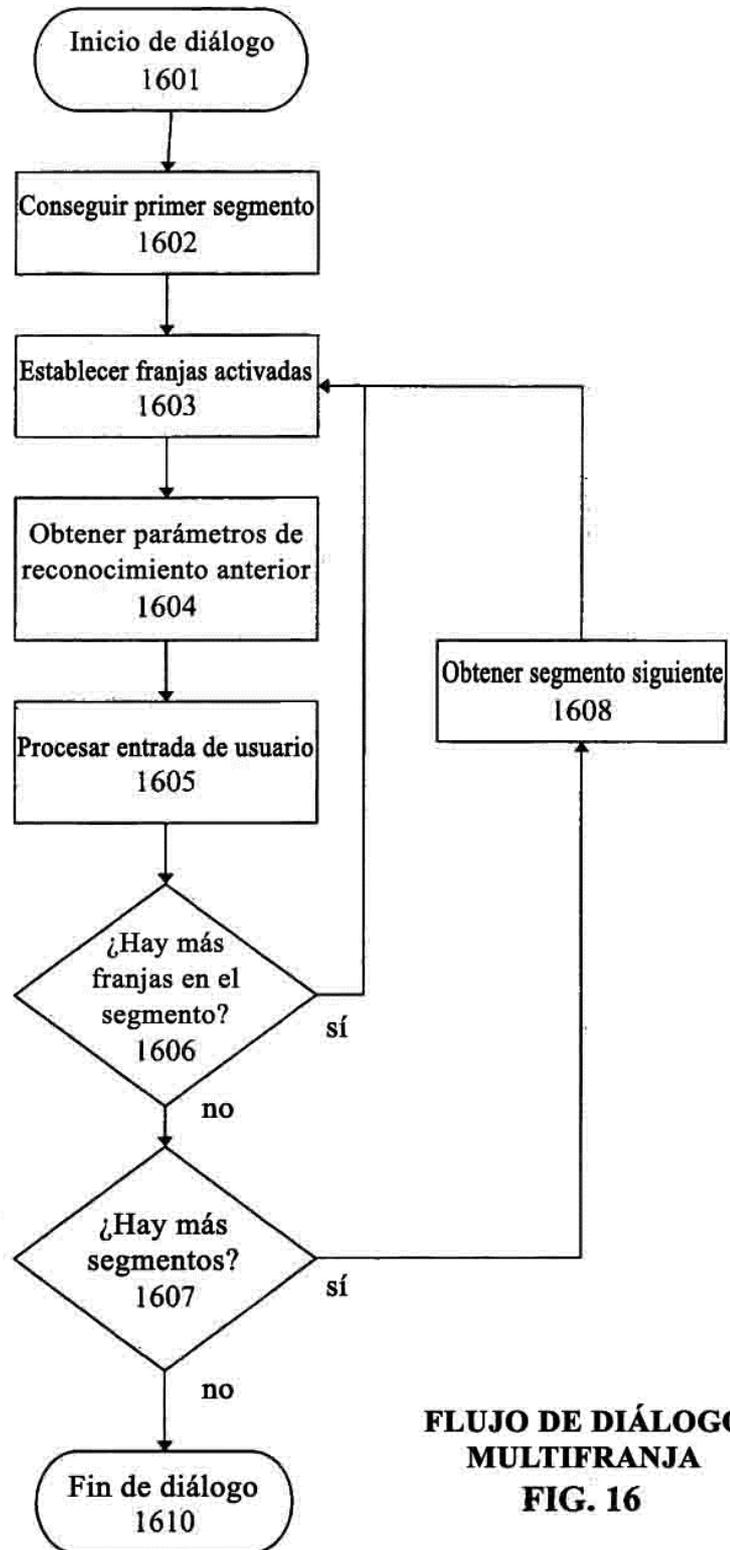
Grupo de franjas	Mensaje guía principal	Mensaje guía de ayuda ....
	<p>¿Cuál es su itinerario de vuelo?</p>	<p>Por ejemplo, puede decir "De Boston a San Francisco mañana a las 6:30 de la tarde"</p>
	<p>¿Desde qué ciudad y a qué ciudad es su vuelo?</p>	<p>Indique por favor desde qué ciudad y qué ciudad va usted a volar. Por ejemplo, puede decir "De Boston a San Francisco"</p>
	<p>¿Cuál es su destino final?</p>	<p>Indique por favor su ciudad de destino final, por ejemplo "Los Ángeles, California"</p>
	<p>¿Cuándo va usted a volar?</p>	<p>Estoy entendiendo la fecha y hora de salida, por ejemplo "16 de agosto hacia las 3 de la tarde"</p>
	<p>¿En qué fecha va a volar?</p>	<p>Puede decir algo como "15 de agosto" o "mañana" o "el próximo lunes"</p>
	<p>¿A qué hora?</p>	<p>Una hora aproximada. Por ejemplo, "hacia las 3 de la tarde"</p>
	<p>¿De la mañana o de la tarde?</p>	<p>No estoy seguro de si la hora es por la mañana o por la tarde. Indique por favor mañana o tarde</p>

GRUPO DE FRANJAS DE RECONOCIMIENTO ANTERIOR

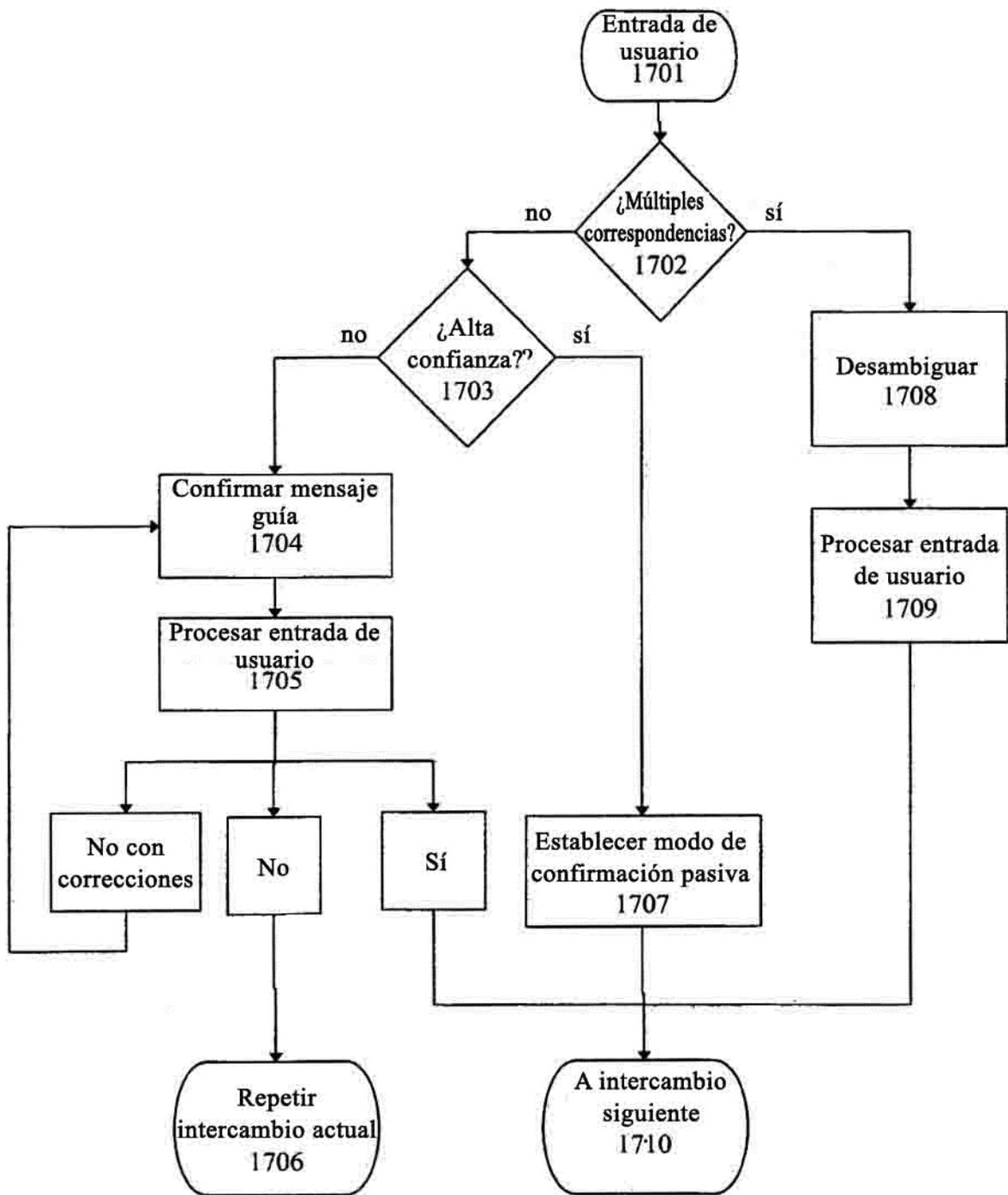
FIG. 14

Grupo de franjas	Mensaje guía para aceptar confirmación	Mensaje guía de encabezamiento de desambiguación
	<p><i>Va usted a volar de &lt;De&gt; a &lt;A&gt; el &lt;Fecha&gt; a las &lt;Hora&gt; &lt;mañana-tarde&gt;</i></p>	<p><i>He encontrado varios itinerarios posibles.</i></p>
	<p><i>Va usted a volar de &lt;De&gt; a &lt;A&gt;</i></p>	<p><i>He encontrado varias combinaciones posibles de ciudades de salida y de destino.</i></p>
	<p><i>Va usted a volar a &lt;A&gt;</i></p>	<p><i>He encontrado varias ciudades de destino posibles.</i></p>
	<p><i>Va usted a volar el &lt;Fecha&gt; a las &lt;Hora&gt; &lt;mañana-tarde&gt;</i></p>	<p><i>He encontrado varias horas de vuelo posibles.</i></p>
	<p><i>Va usted a volar el &lt;Fecha&gt;</i></p>	<p><i>He encontrado varias fechas de vuelo posibles.</i></p>

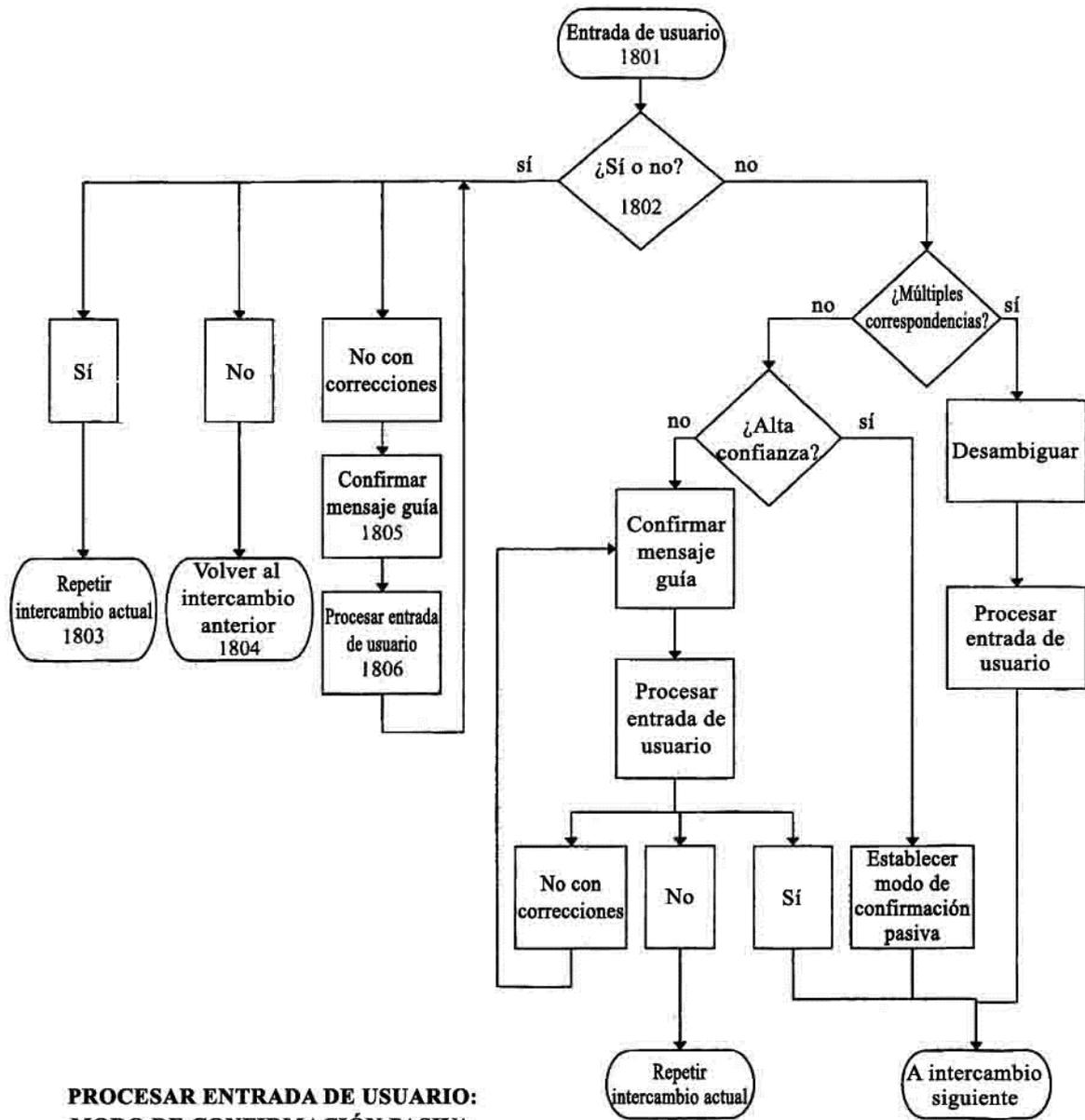
**GRUPO DE FRANJAS DE RECONOCIMIENTO POSTERIOR  
FIG. 15**



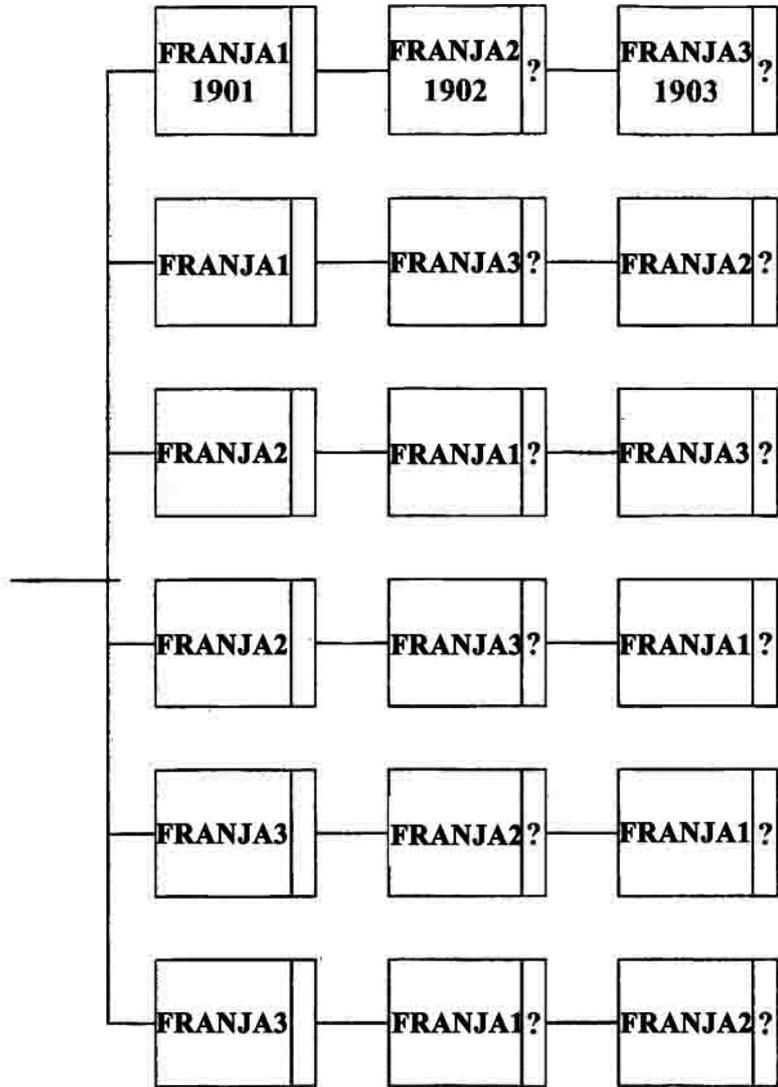
**FLUJO DE DIÁLOGO  
MULTIFRANJA  
FIG. 16**



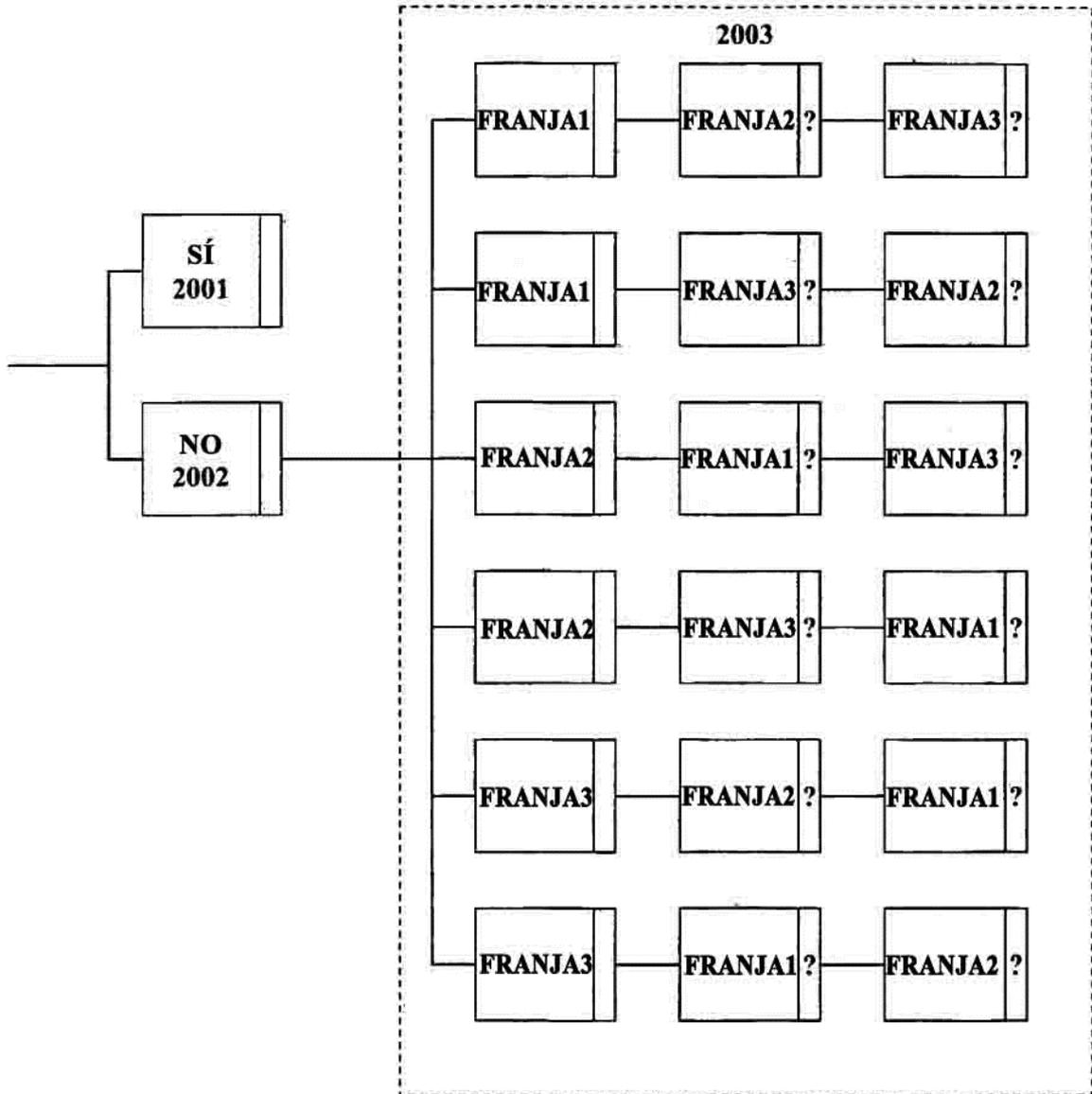
**PROCESAR ENTRADA DE USUARIO:  
MODO NORMAL  
FIG. 17**



**PROCESAR ENTRADA DE USUARIO:  
MODO DE CONFIRMACIÓN PASIVA  
FIG. 18**



**GRAMÁTICA DE TURNO DE USUARIO PRIMARIO  
FIG. 19**



**GRAMÁTICA DE TURNO DE CONFIRMACIÓN**  
**FIG. 20**