



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 603 05 645 T2 2007.05.03**

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 1 473 707 B1**

(51) Int Cl.⁸: **G10L 13/08 (2006.01)**

(21) Deutsches Aktenzeichen: **603 05 645.8**

(96) Europäisches Aktenzeichen: **03 257 090.5**

(96) Europäischer Anmeldetag: **11.11.2003**

(97) Erstveröffentlichung durch das EPA: **03.11.2004**

(97) Veröffentlichungstag

der Patenterteilung beim EPA: **31.05.2006**

(47) Veröffentlichungstag im Patentblatt: **03.05.2007**

(30) Unionspriorität:

2002071306 15.11.2002 KR

(84) Benannte Vertragsstaaten:

DE, FR, GB

(73) Patentinhaber:

**Samsung Electronics Co., Ltd., Suwon, Kyonggi,
KR**

(72) Erfinder:

**Chung, Seung-nyang, Seoul, KR; Cho, no 311-904
3 danji APT, Jeong-mi, Suwon-city Kyungki-do, KR**

(74) Vertreter:

**Patentanwälte Ruff, Wilhelm, Beier, Dauster &
Partner, 70174 Stuttgart**

(54) Bezeichnung: **System und Verfahren zur Text-zu-Sprache Umsetzung mit einer Funktion zur Bereitstellung zusätzlicher Information**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

Beschreibung

[0001] Die vorliegende Erfindung betrifft ein System zur Umsetzung von Text in Sprache und ein Verfahren mit einer Funktion zum Bereitstellen zusätzlicher Information, und insbesondere ein System zur Umsetzung von Text in Sprache und ein Verfahren mit einer Funktion zum Bereitstellen zusätzlicher Information, worin einem Benutzer als zusätzliche Information Wörter bereitgestellt werden, die zu spezifischen Teilen von Sprache gehören oder wovon erwartet wird, dass es für den Benutzer in einem Eingabetext schwierig zu erkennen ist, unter Verwendung von Sprachanalysedaten und Sprachsyntheseergebnisanalysedaten, die in Prozessen der Sprachanalyse und Sprachsynthese eines Systems zur Umsetzung von Text in Sprache (nachfolgend als „TTS“, text-to-speech bezeichnet) erhalten sind, das Text in Sprache umwandelt.

[0002] Bei der Sprachsynthesetechnologie wird, wenn ein Text eingegeben wird, der Text in natürliche, synthetisierte Laute umgewandelt, die wiederum durch Verfahrensweisen der Sprachanalyse des Eingabetextes und seiner Synthese in Sprache ausgegeben werden, die durch das TTS durchgeführt werden.

[0003] Mit Bezug zu [Fig. 1](#) wird eine schematische Konfiguration und ein Verarbeitungsprozess eines allgemeinen TTS durch ein System erklärt, das koreanischen Text in Sprache synthetisiert.

[0004] Zunächst führt eine Vorverarbeitungseinheit **2** einen Vorverarbeitungsprozess zum Analysieren eines eingegebenen Textes unter Verwendung eines Wörterbuchs (Verzeichnisses) von Numeralen/Abkürzungen/Symbolen DB1 durch und ändert dann Zeichen, die keine koreanischen Zeichen sind, in relevante koreanische Zeichen. Die Morphemanalyseeinheit analysiert Morpheme des vorverarbeiteten Satzes unter Verwendung eines Wörterbuchs von Morphemen DB **3** und teilt den Satz gemäß der Morpheme in Teile der Sprache wie Nomen, Adjektiv, Adverb und Partikel.

[0005] Eine syntaktische Analyseeinheit **5** analysiert die Syntax des eingegebenen Satzes. Eine Zeichen/Phonem-Konversionseinheit **7** konvertiert die Zeichen der analysierten Syntax in Phoneme unter Verwendung eines Wörterbuchs spezieller Aussprache DB **6**, das Ausspracheregeldaten zu Symbolen oder speziellen Zeichen speichert.

[0006] Eine Erzeugungseinheit **8** für Sprachsynthesedaten erzeugt einen Rhythmus für das in der Zeichen/Phonem-Konvertierungseinheit **7** konvertierte Phonem; Syntheseeinheiten; Randinformation zu Zeichen, Wörtern und Sätzen; und Dauer der Information zu jedem Stück Sprachdaten. Eine Basisfrequenzsteuereinheit **10** setzt und regelt eine Basisfrequenz der zu synthetisierenden Sprache.

[0007] Ferner führt eine Erzeugungseinheit **11** für synthetische Laute die Sprachsynthese durch Zugriff auf eine Sprachsyntheseeinheit aus, die aus einer Syntheseeinheit DB **12** erhalten ist, die verschiedene synthetisierte Lautdaten speichert, durch die obigen Komponenten erzeugte Sprachsynthesedaten, Information zur Dauer und die Basisfrequenz.

[0008] Gegenstand dieser TTS ist, einem Benutzer zu ermöglichen, dass er die bereitgestellte Textinformation aus den synthetisierten Lauten leicht erkennt. Indessen hat die Sprache eine Zeiteinschränkung, in dem es schwierig ist, eine erneute Rückmeldung der Sprache zu geben, die schon ausgegeben ist, da Sprachinformation mit dem Lauf der Zeit verschwindet. Außerdem ist es unbequem, dass zum Erkennen von Information, die in Form synthetisierter Laute gegeben wird, der Benutzer kontinuierlich seine Aufmerksamkeit auf die Ausgabe der synthetisierten Laute richten muss, und immer versuchen muss, den Inhalt der synthetisierten Laute zu verstehen.

[0009] Indessen ist die Synthese von Text zu Sprache noch nicht perfekt, obwohl Versuche unternommen wurden, natürliche synthetisierte Laute nahe zu einem Eingabetext unter Verwendung von Zeichenerkennung und Synthesedaten in Form einer Datenbank zu erzeugen. Daher kann es sein, dass der Benutzer die von der TTS bereitgestellte Information nicht erkennt oder falsch versteht.

[0010] Deshalb besteht ein Bedarf an einen zusätzlichen Mittel zur reibungslosen Kommunikation durch synthetisierte Laute, die durch ein TTS bereitgestellt werden.

[0011] Zur Lösung der Probleme aus dem Stand der Technik offenbart die koreanischen Patentoffenlegungsschrift Nr. 2002-0011691 mit dem Titel "Graphic representation method of conversation contents and apparatus thereof" ein System, das in der Lage ist, die Effizienz der Konversation zu verbessern, indem gezielt Objekte die in der Konversation enthalten sind, aus einer graphischen Datenbank extrahiert werden, und die Bewegun-

gen, Positionen, Status und dergleichen der extrahierten gezielten Objekte auf einem Bildschirm ausgegeben werden.

[0012] In diesem System zeigt sich der Nachteil, dass eine riesige graphische Datenbank erforderlich ist, um Wörter auszudrücken, die einer Vielzahl von gezielten Objekten entsprechen, die im Alltag verwendet werden, und graphische Information, die jedem Wort entspricht, das eines der gezielten Objekte betrifft, muss in der graphischen Datenbank gesucht und ausgegeben werden.

[0013] Ferner offenbaren die japanische Patentoffenlegungsschrift Nr. 1995-334507 (mit dem Titel "Human body action and speech generation system from text") und die japanische Patentoffenlegungsschrift Nr. 1999-272383 (mit dem Titel "Method and device for generating action synchronized type speech language expression and storage medium storing action synchronized type speech language expression generating program") ein Verfahren, in dem Wörter zur Angabe von Bewegungen aus einem Text extrahiert werden und ein Bewegungsvideo zusammen mit synthetisierten Lauten ausgegeben wird, oder das Bewegungsvideo in Verbindung mit den synthetisierten Lauten werden ausgegeben, wenn Zeichenfolgen in Verbindung mit Bewegungen aus der Sprache erfasst sind.

[0014] Selbst bei diesen Verfahren zeigt sich jedoch der Nachteil, dass eine riesige Datenbank, die das Bewegungsvideo speichert, das die Bewegung zeigt für jeden Text oder jede Zeichenfolge, vorgesehen sein sollte, und wenn ein Text oder eine Zeichenfolge erfasst ist, sollte das relevante Bewegungsvideo aus der Datenbank gesucht und ausgegeben werden.

[0015] Ferner offenbart die koreanische Patentoffenlegungsschrift Nr. 2001-2739 (mit dem Titel "Automatic caption inserting apparatus and method using speech recognition equipment") ein System, worin Schriftdaten durch Erkennen von Sprachsignalen erzeugt werden, die aus einem Tonstreifen eines Programms reproduziert/ausgegeben sind, und es wird veranlasst, dass die Schriftdaten mit dem ursprünglichen Ausgabebetakt der Sprachsignale zusammenfallen, und sie dann ausgegeben werden.

[0016] Da dieses System jedoch nur die Schriftdaten der Sprachsignale anzeigt, die aus dem Tonstreifen reproduziert/ausgegeben sind, ist es kein geeignetes Mittel, das einem Benutzer ermöglicht, die gegebene Information effizienter zu verstehen und zu erkennen.

[0017] Die vorliegende Erfindung stellt ein System zur Umsetzung von Text in Sprache zur Verfügung, das die Funktion hat, zusätzliche Information bereitzustellen.

[0018] Gemäß einem Aspekt der vorliegenden Erfindung wird ein System zur Umsetzung von Text in Sprache zur Verfügung gestellt, umfassend: einen Sprachsynthesemodul zum Analysieren von Textdaten nach Morphen und einer syntaktischen Struktur, Synthetisieren der Textdaten in Sprache unter Verwendung der erhaltenen Sprachsyntheseanalysedaten und Ausgeben synthetisierter Laute; einen Emphasewortwahlmodul zum Auswählen von Wörtern, die zu spezifischen Teilen der Sprache gehören als Emphasewörter aus den Textdaten unter Verwendung der aus dem Sprachsynthesemodul erhaltenen Sprachsyntheseanalysedaten; und einen Anzeigemodul zum Anzeigen der ausgewählten Emphasewörter in Synchronisation mit den synthetisierten Lauten.

[0019] Gemäß einem anderen Aspekt der vorliegenden Erfindung wird ein System zur Umsetzung von Text in Sprache zur Verfügung gestellt, umfassend: einen Informationstypbestimmungsmodul zum Bestimmen des Informationstyps der Textdaten unter Verwendung der vom Sprachsynthesemodul erhaltenen Sprachsyntheseanalysedaten und Erzeugen von Satzmusterinformation; und einen Anzeigemodul zum Umordnen der ausgewählten Emphasewörter nach der erzeugten Satzmusterinformation und Anzeigen der umgeordneten Emphasewörter in Synchronisation mit den synthetisierten Lauten.

[0020] In einer Ausführungsform der vorliegenden Erfindung umfasst das System zur Umsetzung von Text in Sprache ferner einen Strukturierungsmodul zum Strukturieren der ausgewählten Emphasewörter nach einem bestimmten Layoutformat.

[0021] Außerdem beinhalten die Emphasewörter ferner Wörter, die bei denen Übereinstimmungsraten geringer sind als ein bestimmter Schwellenwert und von denen erwartet wird, dass sie für den Benutzer aufgrund von Verzerrung der synthetisierten Laute unter den Wörtern der Textdaten schwierig zu erkennen sind, unter Verwendung der Sprachsyntheseanalysedaten, die aus dem Sprachsynthesemodul erhalten sind, und sie werden unter den ausgewählten Emphasewörtern als Wörter ausgewählt, deren Emphasefrequenzen geringer

sind als ein bestimmter Schwellenwert.

[0022] Gemäß einem anderen Aspekt der vorliegenden Erfindung wird ein Verfahren zur Umsetzung von Text in Sprache zur Verfügung gestellt, umfassend: einen Sprachsyntheseschritt zum Analysieren von Textdaten nach Morphemen und einer syntaktischen Struktur, Synthetisieren der Textdaten in Sprache unter Verwendung erhaltener Sprachsyntheseanalysedaten und Ausgeben synthetisierter Laute; einen Emphasewortauswahlschritt zum Auswählen von Wörtern, die zu spezifischen Teilen von Sprache gehören als Emphasewörter aus den Textdaten unter Verwendung der Sprachsyntheseanalysedaten; und einen Anzeigeschritt zum Anzeigen der ausgewählten Emphasewörter in Synchronisation mit den synthetisierten Lauten.

[0023] Gemäß einem anderen Aspekt der vorliegenden Erfindung wird ein Verfahren zur Umsetzung von Text in Sprache zur Verfügung gestellt, umfassend: einen Satzmusterinformationserzeugungsschritt zum Bestimmen des Informationstyps der Textdaten unter Verwendung der im Sprachsyntheseschritt erhaltenen Sprachsyntheseanalysedaten und Erzeugen von Satzmusterinformation; und worin der Anzeigeschritt ferner zum Umordnen der ausgewählten Emphasewörter nach der erzeugten Satzmusterinformation vorgesehen ist und zum Anzeigen der umgeordneten Emphasewörter in Synchronisation mit den synthetisierten Lauten.

[0024] In einer Ausführungsform der vorliegenden Erfindung umfasst das Verfahren zur Umsetzung von Text in Sprache ferner einen Strukturierungsschritt zum Strukturieren der ausgewählten Emphasewörter nach einem bestimmten Layoutformat.

[0025] Außerdem beinhalten die Emphasewörter ferner Wörter, bei denen Übereinstimmungsraten geringer sind als ein bestimmter Schwellenwert und von denen erwartet wird, dass sie für den Benutzer aufgrund der Verzerrung der synthetisierten Laute schwierig zu erkennen sind, unter Verwendung der Sprachsyntheseanalysedaten, und werden unter den ausgewählten Emphasewörtern als Wörter ausgewählt, deren Emphasefrequenzen geringer sind als ein bestimmter Schwellenwert.

[0026] Die vorliegende Erfindung ermöglicht auf diese Weise reibungslose Kommunikation durch TTS, indem Wörter als Emphasewörter bereitgestellt werden, die zu spezifischen Teilen von Sprache gehören oder von denen erwartet wird, dass sie für einen Benutzer schwierig zu erkennen sind, unter Verwendung von Sprachanalysedaten und Sprachsyntheseergebnisanalysedaten, die im Prozess der Sprachanalyse und Sprachsynthese des TTS erhalten sind.

[0027] Die vorliegende Erfindung verbessert auch die Zuverlässigkeit der TTS durch die Erhöhung der Informationsangabekapazität durch Bereitstellen strukturell angeordneter Emphasewörter zusammen mit synthetisierten Lauten durch die strukturell ausgedrückten Emphasewörter.

[0028] Die obigen und weitere Merkmale der vorliegenden Erfindung werden aus der folgenden Beschreibung bevorzugter Ausführungsformen ersichtlich, die in Verbindung mit den begleitenden Zeichnungen gegeben werden, in denen:

[0029] [Fig. 1](#) ein Diagramm ist, das schematisch eine Konfiguration und den Arbeitsprozess eines herkömmlichen TTS zeigt;

[0030] [Fig. 2](#) ein Blockdiagramm ist, das schematisch eine Konfiguration eines Systems zur Umsetzung von Text in Sprache, gemäß der vorliegenden Erfindung darstellt, das eine Funktion zum Bereitstellen zusätzlicher Information aufweist;

[0031] [Fig. 3](#) ein Fließbild ist, das einen Arbeitsprozess eines Verfahrens zur Umsetzung von Text in Sprache, gemäß einer Ausführungsform der vorliegenden Erfindung darstellt, das eine Funktion zum Bereitstellen zusätzlicher Information aufweist;

[0032] [Fig. 4](#) ein Fließbild ist, das den in [Fig. 3](#) gezeigten Schritt S30 erläutert;

[0033] [Fig. 5](#) ein Fließbild ist, das einen Arbeitsprozess eines Verfahrens zur Umsetzung von Text in Sprache, gemäß einer anderen Ausführungsform der vorliegenden Erfindung darstellt, das eine Funktion zum Bereitstellen zusätzlicher Information aufweist;

[0034] [Fig. 6](#) ein Fließbild ist, das den in [Fig. 5](#) gezeigten Schritt S300 erläutert;

- [0035] [Fig. 7](#) ein Fließbild ist, das den in [Fig. 4](#) gezeigten Schritt S500 erläutert;
- [0036] [Fig. 8](#) eine Ansicht ist, die ein Berechnungsergebnis einer Übereinstimmungsrate gemäß einer anderen Ausführungsform der vorliegenden Erfindung darstellt; und
- [0037] [Fig. 9a](#) bis [Fig. 9c](#) Ansichten sind, die endgültige Zusatzinformation gemäß entsprechender Ausführungsformen der vorliegenden Erfindung zeigen.
- [0038] Nachfolgend werden eine Konfiguration und Arbeitsweise eines Systems zur Umsetzung von Text in Sprache, gemäß der vorliegenden Erfindung, mit einer Funktion zum Bereitstellen zusätzlicher Information ausführlich mit Bezug zu den begleitenden Zeichnungen beschrieben.
- [0039] Mit Bezug zu [Fig. 2](#) umfasst das System zur Umsetzung von Text in Sprache, gemäß einer Ausführungsform der vorliegenden Erfindung hauptsächlich einen Sprachsynthesemodul **100**, einen Emphasewortwahlmodul **300** und einen Anzeigemodul **900**. Eine weitere Ausführungsform der vorliegenden Erfindung weist ferner einen Informationstypbestimmungsmodul **500** und einen Strukturierungsmodul **700** auf.
- [0040] Obwohl eine Historien-DB **310**, eine Domänen-DB **510** und eine Meta-DB **730** wie in [Fig. 2](#) gezeigt, die in den Modulen enthalten sind, in einer Datenbank konstruiert sind (nicht gezeigt), die in einem zusätzlichen Informationserzeugungsgerät gemäß der vorliegenden Erfindung vorgesehen sind, sind sie zur ausführlichen Beschreibung der vorliegenden Erfindung getrennt gezeigt.
- [0041] Der Sprachsynthesemodul **100** analysiert Textdaten ausgehend von Morphem und Syntax, synthetisiert die Eingabetextdaten in Laute unter Verwendung von Sprachanalysedaten und Sprachsyntheseergebnisanalysedaten, die durch die Analyse der Textdaten erhalten sind und gibt die synthetisierten Laute aus. Der Sprachsynthesemodul **100** beinhaltet eine Morphemanalyseeinheit **110**, eine Syntaxanalyseeinheit **130**, eine Sprachsyntheseeinheit **150**, eine Erzeugungseinheit **170** für synthetisierte Laute und einen Lautsprecher **SP 190**.
- [0042] Die Morphemanalyseeinheit **110** analysiert die Morpheme der Eingabetextdaten und bestimmt Teile der Sprache (zum Beispiel Nomen, Pronomen, Partikel, Anhang, Betonung, Adjektiv, Adverb und dergleichen) gemäß der Morpheme. Die Syntaxanalyseeinheit **130** analysiert die Syntax der Eingabetextdaten.
- [0043] Die Sprachsyntheseeinheit **150** führt Synthese von Text in Sprache unter Verwendung der Sprachanalysedaten durch, die in den Morphem- und Syntaxanalyseprozessen von der Morphemanalyseeinheit **110** und der Syntaxanalyseeinheit **130** erhalten sind, und wählt Daten synthetisierter Laute zu entsprechenden Phonemen aus der Syntheseeinheit-DB **12** aus und kombiniert sie.
- [0044] Beim Prozess, in dem die Sprachsyntheseeinheit **150** die entsprechenden Phoneme kombiniert, wird Taktinformation zu den entsprechenden Phonemen erzeugt.
- [0045] Es wird ein Zeitplan für jedes Phonem ausgehend von dieser Taktinformation erzeugt. Deshalb kann der Sprachsynthesemodul **100** durch den erzeugten Zeitplan im Voraus sehen, welches Phonem nach einer bestimmter Zeitspanne (im Allgemeinen auf Basis von 1/1000 sec) ab dem Startpunkt der Sprachsynthese geäußert wird.
- [0046] Das heißt, durch Information eines Startpunkts der Äußerung und gleichzeitiges Betreiben eines Taktgebers, wann die synthetisierten Laute durch den Sprachsynthesemodul **100** ausgegeben werden, können andere Module durch die bei Äußerung des spezifischen Worts (Kombination von Phonemen) bereitgestellte Information, einen Moment abschätzen, wann ein spezifisches Wort geäußert wird.
- [0047] Die Erzeugungseinheit **170** für synthetisierte Laute verarbeitet die Sprachsyntheseergebnisanalysedaten, die von der Sprachsyntheseeinheit **150** erhalten sind, so dass durch den Lautsprecher **190** ausgegeben wird, und gibt sie in Form von synthetisierten Lauten aus.
- [0048] Nachfolgend werden die Sprachanalysedaten, die das Morphem und Syntaxanalysedaten enthalten, die bei den Morphem- und Syntaxanalyseprozessen durch die Morphemanalyseeinheit **110** und die Syntaxanalyseeinheit **130** erhalten sind, und die Sprachsyntheseergebnisanalysedaten, die aus den synthetisierten Lauten gebildet sind, die beim Sprachsyntheseprozess der Sprachsyntheseeinheit **150** erhalten sind, als Sprachsyntheseanalysedaten definiert.

[0049] Der Emphasewortwahlmodul **300** wählt Emphasewörter (zum Beispiel Schlüsselwörter) unter Verwendung von aus dem Sprachsynthesemodul **100** erhaltenen Sprachsyntheseanalysedaten aus den Eingabetextdaten aus und weist eine Historien-DB **310**, eine Emphasewortauswahleinheit **330** und eine Historienverwaltung **350** wie in [Fig. 2](#) gezeigt auf.

[0050] Die Historien-DB **310** speichert Information zu Emphasefrequenzen von Wörtern, die häufig verwendet werden oder unter den aus dem Sprachsynthesemodul **100** erhaltenen Eingabetextdaten empha­siert sind.

[0051] Außerdem speichert sie Information zu Emphasefrequenzen von Wörtern, die häufig verwendet werden oder im Bereich des Informationstyps, der den Eingabetextdaten entspricht, empha­siert sind.

[0052] Die Emphasewortauswahleinheit **330** extrahiert Wörter als Emphasewörter, die zu spezifischen Teilen der Sprache gehören oder von denen erwartet wird, dass sie Verzerrung der synthetisierten Laute aufweisen (d. h. Übereinstimmungs­raten, deren jede aus einer Differenz zwischen einem Ausgabewert, der als synthetisierter Laut erwartet ist, und einem tatsächlichen Ausgabewert berechnet ist), unter Verwendung der vom Sprachsynthesemodul **100** erhaltenen Sprachsyntheseanalysedaten. Außerdem werden die Emphasewörter durch Bezugnahme auf Wörter ausgewählt, die nicht empha­siert werden müssen und von der Historienverwaltung **350** ausgewählt sind.

[0053] Die spezifischen Teile der Sprache sind bestimmte Teile der Sprache, die zum Auswählen der Emphasewörter bezeichnet sind. Wenn die als Emphasewörter ausgewählten Teile der Sprache zum Beispiel ein richtiges Nomen, Lehnwort, eine Zahl oder dergleichen ist, extrahiert die Emphasewortauswahleinheit **330** Wörter entsprechend der bezeichneten Teile der Sprache aus entsprechenden Wörtern, die ausgehend vom Morphem unter Verwendung der Sprachsynthesedaten unterteilt sind.

[0054] Ferner wird die Übereinstimmungsrate des synthetisierten Lauts durch Mittelwertbildung von Übereinstimmungs­raten von Sprachsegmenten unter Verwendung der folgenden Gleichung 1 bestimmt. Es wird angenommen, dass die Verzerrung des synthetisierten Lauts auftreten kann, wenn ein Mittelwert der Übereinstimmungs­raten kleiner ist als ein bestimmter Schwellenwert, und es wird angenommen, dass die Verzerrung des synthetisierten Lauts nur gering oder nicht auftreten kann.

ΣQ (Größe der (Eingabe), |Schätzwert – Istwert|, C)/N, (1)

wo C ein Übereinstimmungswert (Konnektivität) ist und N ein normalisierter Wert (Normalisierung) ist.

[0055] In Gleichung 1 bedeutet die Größe der (Eingabe) die Größe einer Population der ausgewählten Sprachsegmente in der Synthese­einheit-DB, C bedeutet Information zur Verbindung zwischen den Sprachsegmenten und der Schätzwert und der tatsächliche Wert bedeuten einen geschätzten Wert für Länge, Größe und Abstand des Sprachsegments und eines tatsächlichen Werts des ausgewählten Sprachsegments.

[0056] Die Historienverwaltung **350** wählt Wörter aus, deren Emphasefrequenzen den Schwellenwert übersteigen als Wörter, die keine Emphase benötigen, aus Emphasewörtern, die von der Emphasewortauswahleinheit **330** ausgewählt sind, durch Bezugnahme zur Emphasefrequenzinformation, die in der Historien-DB **310** gespeichert ist.

[0057] Der Schwellenwert ist ein Wert, der den Grad angibt, in dem der Benutzer Wörter leicht erkennen kann, da die Wörter häufig benutzt werden oder im Eingabetext empha­siert sind. Zum Beispiel ist dieser Wert auf einen numerischen Wert wie 5-mal gesetzt.

[0058] Der Informationstypbestimmungsmodul **500** bestimmt den Informationstyp der Eingabetextdaten unter Verwendung der aus dem Sprachsynthesemodul **100** erhaltenen Sprachsyntheseanalysedaten und erzeugt Satzmusterinformation. Außerdem weist er eine Domänen-DB **510** auf, eine Semantikanalyseeinheit **530** und eine Erzeugungseinheit **550** für Satzmusterinformation.

[0059] Hier gibt der Informationstyp den Bereich des Typs (nachfolgend als "Domäne" bezeichnet) an, den im Eingabetext gegebene Information darstellt, und die Satzmusterinformation gibt die allgemeine Struktur der vorliegenden Information zum Anzeigen der ausgewählten Emphasewörter als am besten geeignet für den Informationstyp des Eingabetextes an.

[0060] Wenn zum Beispiel ein Text über den Wertpapiermarkt wie "The NASDAQ composite index closed

down 40.30 to 1,356.95" (Der Nasdaq-Index schloss mit einem Minus von 40,30 bei 1.395,95) eingegeben ist, ist der Informationstyp des Eingabetextes der aktuelle Stand der Wertpapiere und die Satzmusterinformation ist ein INDEX-VALUE-Typ, der eine allgemeine Struktur von Nominalphrasen (INDEX) und Zahlen (VALUE) ist, die der vorliegenden Information im aktuellen Status von Wertpapieren ist, was der Informationstyp des Eingabetextes ist.

[0061] Information zu Grammatikregeln, Terminologien und Phrasen für Information, die entsprechend des Informationstyps unterteilt ist, wird als Domäneninformation in der Domänen-DB **510** gespeichert.

[0062] Jede der Grammatikregeln wird dadurch erhalten, dass eine Informationsstruktur jeder Domäne zu Grammatik wird, so dass Teile, die der Information entsprechen, aus einer syntaktischen Struktur des Eingabetextes extrahiert werden können.

[0063] Zum Beispiel stellt die im obigen Beispielsatz verwendete Grammatikregel nur den Preiswert eines Wertpapiers dar, der für den Benutzer von Bedeutung ist, aus "INDEX close (or end) VALUE to VALUE", die eine allgemeine Satzstruktur ist, die im Informationstyp des aktuellen Status von Wertpapieren ist. Die Grammatikregel kann wie folgt definiert werden:

- NP{INDEX}VP{Verb(close)PP{*}PP{to VALUE}} → INDEX VALUE,
- NP{INDEX}VP{Verb(end)PP{*}PP{to VALUE}} → INDEX VALUE.

[0064] Außerdem ist die Terminologie- und Phraseninformation Information zu Wörtern, die häufig verwendet werden oder in spezifischen Domänen emphastriert sind, Phrasen (z. B. "NASDAQ-Index" im obigen Beispielsatz), die als eine semantische Einheit (Block) unterteilt werden können und die Terminologien, die häufig als Abkürzungen in den spezifischen Domänen verwendet werden (z. B. "der NASDAQ-Index" wird im obigen Beispielsatz als "NASDAQ" abgekürzt) und dergleichen.

[0065] Die Semantikanalyseeinheit **530** stellt ein bestimmtes Semantikanalysemittel dar, das zusätzlich vorgesehen ist, wenn Semantikanalyse erforderlich ist, um Semantikinformation zu Textdaten zusätzlich zu den aus dem Sprachsynthesemodul **100** erhaltenen Sprachsyntheseanalysedaten zu erhalten.

[0066] Die Erzeugungseinheit **550** für Satzmusterinformation wählt repräsentative Wörter entsprechend der vorliegenden Information aus den Eingabetextdaten durch Bezugnahme auf die Sprachsyntheseanalysedaten, die vom Sprachsynthesemodul **100** erhalten sind, und die in der Domänen-DB **510** gespeicherte Domäneninformation, bestimmt den Informationstyp und erzeugt die Satzmusterinformation.

[0067] Der Strukturierungsmodul **700** ordnet die ausgewählten Emphasewörter gemäß der von der Erzeugungseinheit **500** für Satzmusterinformation erhaltenen Satzmusterinformation um und passt sie an ein bestimmtes Layoutformat an. Außerdem weist er eine Anpassungseinheit **710** für Satzmusterinformation, eine Meta-DB **730** und eine Informationsstrukturierungseinheit **750** auf, wie es in [Fig. 2](#) gezeigt ist.

[0068] Die Anpassungseinheit **710** für Satzmusterinformation bestimmt, ob die vom Informationstypbestimmungsmodul **500** erzeugte Satzmusterinformation vorhanden ist; wenn die Satzmusterinformation vorhanden ist, passt sie die vom Emphasewortwahlmodul **300** ausgewählten Emphasewörter an die Satzmusterinformation an und gibt sie an die Informationsstrukturierungseinheit **750**; und wenn nicht, gibt sie nur Emphasewörter, die nicht an die Satzmusterinformation angepasst sind, an die Informationsstrukturierungseinheit **750**.

[0069] In der Meta-DB **730** erfolgt Layout (zum Beispiel einer Tabelle) zum strukturellen Anzeigen der ausgewählten Emphasewörter entsprechend dem Informationstyp und dem Inhalt (z. B. ":", ";" usw.) die zusätzlich angezeigt werden sollen.

[0070] Außerdem wird auch Taktinformation zur Metainformation darin gespeichert, um entsprechende Metainformation zusammen mit den synthetisierten Lauten in geeigneter Weise anzuzeigen.

[0071] Die Informationsstrukturierungseinheit **750** extrahiert die Metainformation zu einem relevanten Informationstyp aus der Meta-DB **730** unter Verwendung des Informationstyps und der Emphasewörter für den Eingabetext und die Taktinformation zu den Emphasewörtern, die vom Sprachsynthesemodul **100** erhalten sind; hängt die Emphasewörter und die Taktinformation an die extrahierte Metainformation und gibt sie zum Anzeigemodul **900**.

[0072] Zum Beispiel, wenn es beim Informationstyp des aktuellen Status der Wertpapiere wie im Beispielsatz,

so eingestellt ist, dass INDEX und VALUE, die die vorliegenden Informationen sind, als Layout in Form einer Tabelle angezeigt werden, werden sie mit der Taktinformation (SYNC= "12345", SYNC="12348") für die INDEX-Information und die VALUE-Information, die aus dem Sprachsynthesemodul **100** erhalten ist, versehen.

[0073] Die zusammen mit der Taktinformation strukturierten Emphasewörter im durch diese Vorgehensweise bezeichneten Layoutformat sind wie folgt:

```
<INDEXVALUE ITEM = "1">
  <INDEX SYNC = "12345">INDEX(NASDAQ)</INDEX>
  <VALUE SYNC = "12348">VALUE(1,356.95)</VALUE>
</INDEXVALUE>
```

[0074] Der Anzeigemodul **900** synchronisiert die strukturierten Emphasewörter mit den synthetisierten Lauten entsprechend der Taktinformation und zeigt sie an. Der Anzeigemodul **900** weist eine Synchronisationseinheit **910**, eine Videosignalverarbeitungseinheit **930** und eine Anzeigeeinheit **950** auf, wie in [Fig. 2](#) gezeigt.

[0075] Die Synchronisationseinheit **910** extrahiert entsprechende Taktinformation zur Metainformation und den Emphasewörtern und synchronisiert die durch den Lautsprecher **190** des Sprachsynthesemoduls **100** ausgegebenen synthetisierten Laute mit den Emphasewörtern und der Metainformation, so dass sie richtig angezeigt werden können.

[0076] Die Videosignalverarbeitungseinheit **930** verarbeitet die strukturierten Emphasewörter entsprechend der von der Synchronisationseinheit **910** erhaltenen Taktinformation in Videosignale, so dass sie zur Anzeigeeinheit **950** ausgegeben werden können.

[0077] Die Anzeigeeinheit **950** zeigt die Emphasewörter entsprechend der von der Videosignalverarbeitungseinheit **930** ausgegebenen Anzeigeeinheit **950** sichtbar an.

[0078] Zum Beispiel wird der vom Strukturierungsmodul **700** ausgegebene strukturierte Beispielsatz von der Anzeigeeinheit **950** wie folgt ausgegeben:

NASDAQ	1.356,95
--------	----------

[0079] Nachfolgend wird ein Verfahren zur Umsetzung von Text in Sprache mit der Funktion zur Bereitstellung zusätzlicher Information gemäß der vorliegenden Erfindung ausführlich mit Bezug zu den begleitenden Zeichnungen beschrieben.

[0080] [Fig. 3](#) ist ein Fließbild, das einen Arbeitsprozess des Verfahrens zur Umsetzung von Text in Sprache, gemäß einer Ausführungsform der vorliegenden Erfindung darstellt, das eine Funktion zum Bereitstellen zusätzlicher Information aufweist.

[0081] Zunächst führt der Sprachsynthesemodul **100** die Morphem- und Syntexanalyseprozesse für den Eingabetext durch die Morphemanalyseeinheit **110** und die Syntexanalyseeinheit **130** durch und synthetisiert die Eingabetextdaten in Sprache, indem auf Sprachsyntheseanalysedaten Bezug genommen wird, die durch die Morphem- und Syntexanalyseprozesse erhalten sind (S10).

[0082] Wenn der Sprachsynthesemodul **100** die synthetisierten Laute erzeugt, wählt die Emphasewortauswahleinheit **330** des Emphasewortauswahlmoduls **300** als Emphasewörter Wörter aus, von denen erwartet wird, dass sie für den Benutzer schwierig zu erkennen sind oder zu speziellen Teilen der Sprache gehören, unter Verwendung der vom Sprachsynthesemodul **100** erhaltenen Sprachsyntheseanalysedaten (S30).

[0083] Wenn die Emphasewortauswahleinheit **330** die Emphasewörter auswählt, werden die ausgewählten Emphasewörter und die Taktinformation, die vom Sprachsynthesemodul **100** erhalten sind, zu ihrer Synchronisation verwendet (S50).

[0084] Der Anzeigemodul **900** extrahiert Taktinformation aus den Emphasewörtern, die mit der Taktinformation strukturiert sind, synchronisiert sie mit den vom Lautsprecher **190** des Sprachsynthesemoduls **100** ausgegebenen synthetischen Lauten und zeigt sie auf der Anzeigeeinheit **950** an (S90).

[0085] Außerdem werden die ausgewählten Emphasewörter durch Extrahieren der Metainformation, die dem bestimmten Layoutformat entspricht, aus der Meta-DB **730** strukturiert und die Emphasewörter an die extrahierte Metainformation angepasst (S70).

[0086] [Fig. 4](#) zeigt den Schritt zum Auswählen der Emphasewörter (S30) ausführlicher. Wie in der Figur gezeigt ist, extrahiert die Emphasewortauswahleinheit **330** die vom Sprachsynthesemodul **100** erhaltenen Sprachsyntheseanalysedaten (S31).

[0087] Dann wird bestimmt, ob der Teil der Sprache jedes Worts, das ausgehend vom Morphem entsprechend dem Morphemanalyseprozess, der in der Morphemanalyseeinheit **110** des Sprachsynthesemoduls **100** durchgeführt ist, unterteilt ist, zum spezifischen Teil der Sprache gehört, unter Verwendung der extrahierten Sprachsyntheseanalysedaten, und ein Wort, das dem bezeichneten spezifischen Teil der Sprache entspricht, wird als Emphasewort ausgewählt (S32).

[0088] Außerdem werden die Übereinstimmungsraten der synthetisierten Laute von Wörtern unter Verwendung der extrahierten Sprachsyntheseanalysedaten mittels der Emphasewörter geprüft, um Wörter zu erhalten, von denen angenommen wird, dass sie für den Benutzer schwierig zu erkennen sind (S33). Als Ergebnis der Prüfung der Übereinstimmungsraten der synthetisierten Laute werden Wörter extrahiert, von denen angenommen wird, dass sie Verzerrung der synthetisierten Laute aufweisen und als Emphasewörter ausgewählt (S34).

[0089] Im Falle der Prüfung der Übereinstimmungsraten der synthetisierten Laute wird jede der Übereinstimmungsraten aus der Differenz zwischen dem Ausgabewert (Schätzwert) des synthetisierten Lauts, der für jedes Sprachsegment jedes Worts aus den extrahierten Sprachsyntheseanalysedaten abgeschätzt ist, und dem vorliegenden Ausgabewert (effektiver Wert) des synthetisierten Lauts unter Verwendung der Gleichung 1 berechnet. Es wird ein Wort gesucht, dessen Mittelwert der berechneten Übereinstimmungsraten kleiner ist als der Schwellenwert.

[0090] Der Schwellenwert gibt einen Mittelwert von Übereinstimmungsraten eines synthetisierten Lauts an, den der Benutzer nicht erkennen kann und wird als Zahlenwert wie 50 % angesetzt.

[0091] Ferner wählt die Emphasewortauswahleinheit **330** zum Auswählen von Wörtern, die der Benutzer leicht erkennen kann, aus den durch die obigen Prozesse ausgewählten Emphasewörtern als Wörter, bei denen keine Emphase notwendig ist, Wörter, bei denen keine Emphase notwendig ist, aus den extrahierten Emphasewörtern durch die Historienverwaltung **350** aus (S35).

[0092] Das heißt, die Historienverwaltung **350** wählt Wörter aus, bei denen die Emphasfrequenzen höher sind als der Schwellenwert und die Möglichkeit, dass der Benutzer sie nicht erkennen kann, bei den durch die Emphasewortauswahleinheit **330** ausgewählten Wörtern gering ist, durch Bezugnahme zur in der Historien-DB **310** gespeicherten Emphasfrequenzinformation, die vom Sprachsynthesemodul **100** erhalten ist.

[0093] Die Emphasewortauswahleinheit **330** wählt Wörter aus, die zu spezifischen Teilen der Sprache gehören und von denen erwartet wird, dass sie für den Benutzer schwer zu erkennen sind, aus dem Eingabetext durch den Prozess der Auswahl von Wörtern, bei denen keine Emphase notwendig ist, mit der Historienverwaltung **350** (S36).

[0094] [Fig. 5](#) zeigt einen Spracherzeugungsprozess in einem Verfahren zur Umsetzung von Text in Sprache mit einer Funktion zur Bereitstellung zusätzlicher Information gemäß einer anderen Ausführungsform der vorliegenden Erfindung. Die Ausführungsform von [Fig. 5](#) wird wiederum mit Bezug zu den [Fig. 3](#) und [Fig. 4](#) beschrieben.

[0095] Zunächst wird durch den Sprachsynthesemodul **100** eingegebener Text in Sprache umgewandelt (S100, siehe Schritt S10 in [Fig. 3](#)) und die Emphasewortauswahleinheit **330** wählt Emphasewörter unter Verwendung der Sprachsyntheseanalysedaten, die vom Sprachsynthesemodul **100** erhalten sind (S200, siehe Schritt S30 in den [Fig. 3](#) und [Fig. 4](#)).

[0096] Ferner bestimmt die Erzeugungseinheit **550** für Satzmusterinformation des Informationstypbestimmungsmoduls **500** den Informationstyp des Eingabetextes unter Verwendung der Sprachsyntheseanalysedaten, die vom Sprachsynthesemodul **100** erhalten sind und der Domäneninformation, die von der Domänen-DB **530** extrahiert ist und erzeugt die Satzmusterinformation (S300).

[0097] Dann bestimmt die Anpassungseinheit **710** für Satzmusterinformation der Strukturierungseinheit **700** die Möglichkeit der Anwendung der Satzmusterinformation durch Bestimmen, ob die Satzmusterinformation, für die die ausgewählten Emphasewörter geeignet sind, vom Informationstypbestimmungsmodul **500** erzeugt sind (S400).

[0098] Wenn bestimmt ist, dass die Satzmusterinformation anwendbar ist, wird Umordnung vorgenommen, indem die ausgewählten Emphasewörter an die Satzmusterinformation angepasst werden (S500).

[0099] Dann werden die Emphasewörter, die an das Satzmuster angepasst sind oder nicht mit der vom Sprachsynthesemodul **100** erhaltenen Taktinformation synchronisiert (S600, siehe Schritt S50 in [Fig. 3](#)).

[0100] Der Anzeigemodul **900** extrahiert die Taktinformation aus den Emphasewörtern, die mit der Taktinformation strukturiert sind, synchronisiert sie in geeigneter Weise mit den synthetisierten Lauten, die vom Lautsprecher **190** des Sprachsynthesemoduls **100** ausgegeben werden und zeigt sie auf der Anzeigeeinheit **950** an (S800, siehe Schritt S90 in [Fig. 3](#)).

[0101] Zusätzlich extrahiert die Informationsstrukturierungseinheit **750** des Strukturierungsmoduls **700** Meta-information des relevanten Informationstyps aus der Metainformation-DB **730** und strukturalisiert die Emphasewörter, die an die Satzmusterinformation angepasst sind oder nicht im bestimmten Layoutformat (S700, siehe Schritt S70 in [Fig. 3](#)).

[0102] [Fig. 6](#) zeigt spezifisch Schritt S30 zum Bestimmen des Informationstyps und Erzeugen der Satzmusterinformation in [Fig. 5](#). Der Schritt wird ausführlich als Beispiel mit Bezug zu den Figuren beschrieben.

[0103] Zunächst extrahiert die Erzeugungseinheit **550** für Satzmusterinformation des Informationstypbestimmungsmoduls **500** die Sprachsyntheseanalysedaten aus dem Sprachsynthesemodul **100**, und wenn die Information zur semantischen Struktur des Eingabetextes zusätzlich erforderlich ist, analysiert sie die semantische Struktur des Textes durch die Semantikanalyseeinheit **530** und extrahiert die Bedeutungsstrukturinformation des Eingabetextes (S301).

[0104] Dann werden entsprechende Wörter des Eingabetextes ausgehend von den vorliegenden Semantikeinheiten mit Bezugnahme zu den extrahierten Sprachsyntheseanalysedaten, die Semantikstrukturinformation und die Domänen-DB **510** unterteilt (S302).

[0105] Nach Unterteilen des Eingabetextes ausgehend von den Semantikeinheiten (Blöcke) werden die repräsentativen Bedeutungen zum Angeben unterteilter Semantikeinheiten bestimmt und entsprechende Semantikeinheiten mit der bestimmten Semantikinformation versehen (S303) und repräsentative Wörter der entsprechenden Semantikeinheiten werden durch Bezugnahme zur Domänen-DB **510** ausgewählt (S304).

[0106] Wenn zum Beispiel im obigen Beispielsatz entsprechend dem Informationstyp des aktuellen Status von Wertpapieren, die Semantikeinheiten in `"/The NASDAQ composite index/close/down/40.30/to/1,356.95/"` sind, ist die Semantikinformation, d. h. Information zur Bezeichnung der entsprechenden semantischen Einheiten wie folgt definiert:

- The NASDAQ composite index: INDEX,
- close: close,
- down: down,
- to: to,
- Zahlenklasse (40.30, 1,356.95): VALUE (WERT).

[0107] Wenn die oben definierte Semantikinformation an den Eingabetext angehängt wird, der ausgehend von den Semantikeinheiten unterteilt ist, wird das Folgende gebildet.
/INDEX/close/down/VALUE/to/VALUE.

[0108] Wenn außerdem die repräsentativen Wörter der entsprechenden Semantikeinheiten aus dem Eingabetext ausgewählt werden, der ausgehend von den Semantikeinheiten unterteilt ist, durch Bezugnahme zur Terminologie und Phraseninformation, die in der Domänen-DB **510** gespeichert ist, wird es wie folgt bestimmt:
/NASDAQ/close/down/40.30/to/1,356.951.

[0109] Dem Benutzer als tatsächliche Information bereitzustellende Wörter werden aus den repräsentativen Wörter durch solche Prozesse ausgewählt.

[0110] Nach Auswählen der repräsentativen Wörter extrahiert die Satzmusterinformationserzeugungseinheit **550** die auf die Syntax- und Semantikstruktur des Eingabetexts anwendbare Grammatikregel aus der Domänen-DB **510** und wählt den Informationstyp und die repräsentativen Wörter aus, die als die tatsächliche Information durch die extrahierte Grammatikregel ausgedrückt werden sollen (S305).

[0111] Wenn zum Beispiel unter Bezugnahme auf den Informationstypbestimmungsprozess für den obigen Beispielsatz in der Beschreibung der Grammatikregel, die zuvor in der Domänen-DB **510** gespeichert ist, die Syntaxstruktur des als "NP{INDEX}VP{Verb(close) PP{*}PP{to VALUE}} → INDEX VALUE" eingegebenen Textes der als Grammatikregel des bestimmten Informationstyps bereitgestellten Grammatik entspricht, führt Anpassung des ausgehend von den Semantikeinheiten unterteilten Textes an die erfasste Grammatikregel zum Folgenden:

INFO[The NASDAQ composite index/INDEX] closed town 40.30 to INFO[1,356.95/VALUE].

[0112] Auf diese Weise wird der Informationstyp des Eingabetexts im Prozess der Anwendung der Grammatikregel bestimmt und die repräsentativen Wörter [(INDEX, VALUE)], die als die tatsächliche Information ausgedrückt werden sollen, werden ausgewählt.

[0113] Wenn der Informationstyp bestimmt ist und die repräsentativen Wörter, die als die tatsächliche Information ausgedrückt werden sollen, ausgewählt sind, wird die Satzmusterinformation zum Anzeigen der ausgewählten repräsentativen Wörter passend zum bestimmten Informationstyp erzeugt (S306).

[0114] Zum Beispiel ist die im obigen Beispielsatz erzeugte Satzmusterinformation vom Typ "INDEX WERT".

[0115] [Fig. 7](#) zeigt speziell den Schritt S500 zum Anwenden der Satzmusterinformation in [Fig. 5](#). Der Prozess wird ausführlich als Beispiel mit Bezug zu den Figuren beschrieben.

[0116] Zunächst wird zum Bestimmen, ob die vom Emphasewortauswahlmodul **300** ausgewählten Emphasewörter für die erzeugte Satzmusterinformation geeignet sind, bestimmt, ob die ausgewählten Emphasewörter in den repräsentativen Wörtern enthalten sind, die als die tatsächliche Information ausgedrückt werden sollen, die aus der Satzmusterinformation ausgewählt sind, die von der Satzmusterinformationserzeugungseinheit **550** erzeugt ist (S501).

[0117] Wenn bestimmt ist, dass die ausgewählten Emphasewörter nicht in den repräsentativen Wörtern enthalten sind, werden die ausgewählten Emphasewörter gemäß der Syntaxstruktur des im Prozess zur Erzeugen der Satzmusterinformation (S502) bestimmten Informationstyps umgeordnet, und wenn nicht, werden die Emphasewörter durch Anhängen der Emphasewörter an die relevanten repräsentativen Wörter in der Satzmusterinformation umgeordnet (S503).

[0118] Ausführungsformen, in denen das System zur Umsetzung von Text in Sprache und das Verfahren mit der Funktion zum Bereitstellen zusätzlicher Information gemäß der vorliegenden Erfindung durch ein mobiles Endgerät implementiert sind, werden mit Bezug zu den begleitenden Zeichnungen beschrieben.

[0119] Nachfolgend werden bevorzugte Ausführungsformen der vorliegenden Erfindung mit Bezug zu Prozessen zum Erfassen und Anzeigen von Emphasewörtern, Umordnen der erfassten Emphasewörter gemäß der Syntaxmusterinformation und dann ihre Anzeige und Anwenden der erfassten Emphasewörter auf die Syntaxmusterinformation und dann ihre Organisation mit Metainformation und Anzeige beschrieben.

[0120] Außerdem können Prozesse zur Interpretation von Morphem/Struktur und Erfassung eines Emphaseworts auf verschiedene linguistische Bereiche angewendet werden, und nachfolgend werden Koreanisch und Englisch verwendet.

Ausführungsform 1

[0121] Es wird ein Beispiel erläutert, in dem die Emphasewörter durch den Emphasewortauswahlmodul **300** ausgewählt werden und nur ausgewählte Emphasewörter dann angezeigt werden, wenn der folgende Text eingegeben ist:

“GE 백색 가전은 양문여닫이 냉장고인 ‘GE 프로파일 아티카’를 출시한
다고 8월 9일 밝혔다.”

[GE bæksæk gadjənen yaŋmun yədadjɪ næŋdʒəŋɡoin ‘GE Profile Artica’rel
tfulsihandago 8wol 9il balkyətɔda]

[0122] Dies bedeutet "GE Appliances announced on Aug. 9 that it would present the side-by-side refrigerator, 'GE Profile Artica' " (GE Appliances kündigte am 9. Aug. die Vorstellung des Kühlgeräts 'GE Profile Artica' an).

[0123] Wenn ein solcher Text eingegeben wird, teilt der Sprachsynthesemodul **100** den Eingabetext derart in Sprachteile, dass das Nomen, das Adjektiv, das Adverb und der Partikel gemäß dem Morphem durch die Morphemanalyseeinheit **110** unterteilt werden, so dass die Sprachsynthese des Eingabetexts vorgenommen wird. Das Ergebnis ist wie folgt:

"GE/Fremdwort + 백색 (bæksæk)/Nomen + 가전 (gadjən)/Nomen +
은(en)/Partikel + 양문여닫이(yaŋmunyədadjɪ)/Nomen +
냉장고(næŋdʒəŋɡo)/Nomen + 인(in)/Prädikat + GE/Fremdwort +
프로파일 (Profile)/Nomen + 아티카 (Artica)/Eigename +
를 (rel)/Partikel + 출시한다(tfulsihanda)/Prädikat +
고(go)/Verknüpfungsendung + 8/Numerale + 월(wol)/Nomen +
9/Numerale + 일(il)/Nomen + 밝혔(balkyət)/Prädikat +
다(da)/Endung."

[0124] Nachdem der Satz gemäß dem Morphem durch die Morphemanalyseeinheit **110** analysiert ist, werden die Sprachsyntheseanalysedaten durch die Prozesse zum Analysieren der Satzstruktur der Eingabetextdaten in der Satzstrukturanalyseeinheit **130**, Bezugnahme zur analysierten Satzstruktur und Synthetisieren der Sprache in der Sprachsyntheseinheit **150** erzeugt.

[0125] Die Emphasewortauswahleinheit **330** des Emphasewortauswahlmoduls **300** extrahiert Wörter, die zu den bestimmten spezifischen Teilen der Sprache gehören, aus den Wörtern, die gemäß dem Morphem in den Eingabetextdaten unterteilt sind, unter Verwendung der Sprachsyntheseanalysedaten, die vom Sprachsynthesemodul **100** erhalten sind.

[0126] Wenn in der vorliegenden Ausführungsform der Eigename, das Lehnwort und das Numerale als spezifischer Teil der Sprache bezeichnet sind, extrahiert die Emphasewortauswahleinheit **330** 'GE/아티카/8월9일' vom Eingabetext als Wörter, die zu den bestimmten spezifischen Teilen der Sprache gehören.

[0127] Wenn außerdem Wörter, von denen erwartet wird, dass sie für den Benutzer schwierig zu erkennen sind, als Emphasewörter auszuwählen sind, erfasst die Emphasewortauswahleinheit **330** die Übereinstimmungsrate der synthetisierten Laute der Wörter in den Eingabetextdaten gemäß Gleichung 1.

[0128] Wenn dann die Übereinstimmungsrate des Worts "양문여닫이" zu 20% errechnet ist, wie in [Fig. 8](#) gezeigt, wird das Wort "양문여닫이" als ein Wort erfasst, von dem angenommen wird, dass es Verzerrung des synthetisierten Lauts aufweist, da die berechnete Übereinstimmungsrate geringer ist als der Schwellenwert in einem Fall, wo der eingestellte Schwellenwert 50 % beträgt.

[0129] Durch die Prozesse werden die Wörter "GE/아티카/8월9일/양문여닫이" als die Emphasewörter erfasst die zu den spezifischen Teilen der Sprache gehören und von denen erwartet wird, dass sie Verzerrung der syn-

thetisierten Laute aufweisen.

[0130] Wenn außerdem die Wörter, die im Eingabetext häufig verwendet werden und deren Emphasfrequenzen höher sind als der bestimmte Schwellenwert, aus den gewählten Emphasewörtern als Wörter auszuwählen sind, die keine Emphase benötigen, wählt die Emphasewortauswahleinheit **330** unter den von der Historienverwaltung **350** ausgewählten Emphasewörtern Wörter aus, deren Emphasfrequenzen höher sind als der Schwellenwert.

[0131] Wenn in der Ausführungsform alle ausgewählten Emphasewörter Emphasfrequenzen unter dem Schwellenwert aufweisen, werden endgültige Emphasewörter als die Wörter “GE/아타카/8월9일/강릉여공회” ausgewählt.

[0132] Die Strukturierungsmodul **700** strukturiert die ausgewählten Emphasewörter zusammen mit der Taktinformation, die vom Sprachsynthesemodul **100** erhalten ist. Der Anzeigemodul **900** extrahiert die Taktinformation aus den strukturierten Emphasewörtern und zeigt die Emphasewörter auf der Anzeigeeinheit **950** zusammen mit den synthetisierten Lauten an, die vom Sprachsynthesemodul **100** ausgegeben sind.

[0133] Die auf der Anzeigeeinheit **950** angezeigten Emphasewörter sind in [Fig. 9a](#) gezeigt.

[0134] Darüber hinaus können die ausgewählten Emphasewörter gemäß dem bestimmten Layoutformat angezeigt werden, das aus der Meta-DB **730** extrahiert ist.

Ausführungsform 2

[0135] Es wird ein weiteres Beispiel erläutert, wo die Emphasewörter durch den Emphasewortauswahlmodul **300** ausgewählt und die ausgewählten Emphasewörter ungeordnet und gemäß der Satzmusterinformation angezeigt werden, wenn der folgende Text eingegeben ist: "The whole country will be fine but in the Yongdong district it will become partly cloudy." (Im gesamten Land ist es heiter, aber im Distrikt Yongdong wird es teilweise wolkig).

[0136] Nachfolgend wird angenommen, dass die ausgewählten Emphasewörter den repräsentativen Wörtern der tatsächlichen Information entsprechen, die im Prozess zur Bestimmung des Informationstyps ausgewählt sind. Daher wird die Beschreibung des Prozesses zum Auswählen der Emphasewörter ausgelassen und nur der Prozess zum Anzeigen der Emphasewörter gemäß der Satzmusterinformation wird beschrieben.

[0137] Zunächst unterteilt der Informationstypbestimmungsmodul **500** die Wörter des Eingabetexts ausgehend von ihren vorliegenden Semantikeinheiten unter Bezugnahme auf die Sprachsyntheseanalysedaten, die vom Sprachsynthesemodul **100** erhalten sind und die Domäneninformation, die von der Domänen-DB **510** extrahiert ist. Das Ergebnis wird wie folgt ausgedrückt:
"The whole country/will be/fine/but/in/the Yongdong district/it/will become/partly cloudy/".

[0138] Der Eingabetext wird ausgehend von den vorliegenden Semantikeinheiten geteilt und die repräsentativen Bedeutungen werden dann für die unterteilten Semantikeinheiten bestimmt, so dass die bestimmten repräsentativen Bedeutungen an die entsprechenden Semantikeinheiten angefügt werden. Das Ergebnis mit der angehängten repräsentativen Bedeutung wird wie folgt ausgedrückt:
"/REGION/will be/FINE/but/in/REGION/it/will become/CLOUDY/".

[0139] Wenn außerdem die repräsentativen Wörter der entsprechenden Semantikeinheiten aus dem Eingabetext, der gemäß der Semantikeinheiten geteilt ist, mit Bezugnahme zur Information zu den Terminologien und Phrasen ausgewählt werden, die in der Domänen-DB **510** gespeichert sind, kann das Ergebnis auch wie folgt ausgedrückt werden:
"/whole country/be/fine/but/in/Youngdong/it/become/partly cloudy/".

[0140] Wörter, die dem Benutzer als tatsächliche Information bereitgestellt werden, werden aus den Wörtern ausgewählt, die durch den obigen Prozess ausgewählt sind. Die Satzmusterinformationserzeugungseinheit **550** extrahiert die Grammatikregel, die auf die Syntax- und Semantikstruktur der Textdaten anwendbar ist, aus der Domänen-DB **510**.

[0141] Wenn die folgende Grammatikregel, die auf den in diesem Beispiel angeführten Text anwendbar ist, vom Informationstyp des Wetterberichts auf die selbe Weise wie in der folgenden Regel extrahiert wird, wird

der Informationstyp des Eingabetexts als Wettervorhersage bestimmt.

- NP{REGION}VP{be FINE} → REGION FINE

- PP{in NP{REGION}}NP{it}VP{become CLOUDY} → REGION CLOUDY

[0142] Wenn der Informationstyp bestimmt ist, werden die Eingabetextdaten auf die extrahierte Grammatikregel aufgegeben. Das Ergebnis der Anwendung der Grammatikregel wird wie folgt ausgedrückt:

"INFO[THE whole country/REGION] will be INFO[fine/FINE] but in INFO [the YONGDONG district/REGION] it will become INFO[partily cloudy/CLOUDY]."

[0143] Wie oben beschrieben wird der Informationstyp des Eingabetexts im Prozess zum Anwenden der Grammatikregel bestimmt, und die repräsentativen Wörter (d. h. The whole country/REGION, fine/FINE, the Yongdong district/REGION, partly cloudy/CLOUDY), die als die tatsächliche Information auszudrücken sind, werden ausgewählt.

[0144] Wenn der Informationstyp bestimmt ist und die als die tatsächliche Information auszudrückenden repräsentativen Wörter ausgewählt sind, wird das Satzmuster zum Anzeigen der ausgewählten repräsentativen Wörter auf die am besten geeignete Weise für den bestimmten Informationstyp erzeugt.

[0145] Zum Beispiel ist die aus dem Text erzeugte Satzmusterinformation vom Typ "REGION WETTER".

[0146] Wenn die Satzmusterinformation durch den obigen Prozess erzeugt ist, ordnet die Satzmusterinformationsanpassungseinheit **910** die ausgewählten Emphasewörter gemäß der erzeugten Satzmusterinformation um.

[0147] Wenn in der Ausführungsform die ausgewählten Emphasewörter den Wörtern entsprechen, die aus der Satzmusterinformation als die repräsentativen Wörter ausgewählt sind, die als die tatsächliche Information ausdrücken sind, werden die Emphasewörter und die Taktinformation der entsprechenden Emphasewörter, die aus dem Sprachsynthesemodul **100** erhalten sind, an die Satzmusterinformation angehängt, um die Emphasewörter zu strukturieren.

[0148] Die strukturierten Emphasewörter werden wie folgt ausgedrückt:

<REGION WEATHER ITEM = "3">

>REGION VALUE = "0" SYNC = "1035"> the whole country</REGION>

<WEATHER EVAL = "CLOUD" SYNC="1497">fine<WEATHER>

.

.

</REGIONWEATHER>

[0149] Der Anzeigemodul **900** zeigt die strukturierten Emphasewörter zusammen mit den synthetisierten Lauten in einem Zustand an, wo sie entsprechend der Taktinformation miteinander synchronisiert sind.

[0150] Das Anzeigeergebnis ist in [Fig. 9b](#) gezeigt.

Ausführungsform 3

[0151] Es wird ein weiteres Beispiel erläutert, wo die Emphasewörter vom Emphasewortauswahlmodul **300** ausgewählt werden und die ausgewählten Emphasewörter strukturiert und zusammen mit der Metainformation gemäß der Satzmusterinformation angezeigt werden, wenn der folgende Text eingegeben ist:

"Today, the Nasdaq composite index closed down 0.57 to 1,760.54 and the Dow Jones industrial average finished up 31.39 to 9397.51 ". (Heute schloss der Nasdaq-Index mit einem Minus von 0,57 bei 1.760,54 und der Dow Jones endete mit einem Plus von 31,39 bei 9397,51)

[0152] Nachfolgend wird angenommen, dass die ausgewählten Emphasewörter den repräsentativen Wörtern der tatsächlichen Information entsprechen, die im Prozess zur Bestimmung des Informationstyps ausgewählt ist. Daher wird die Beschreibung des Prozesses zum Auswählen der Emphasewörter ausgelassen und nur der Prozess zum Anzeigen der Emphasewörter gemäß der Satzmusterinformation beschrieben.

[0153] Der Sprachsynthesemodul **100** analysiert den Eingabetext gemäß der Morphem- und Semantikstruktur und synthetisiert den analysierten Text in Sprache.

[0154] Der Emphasewortauswahlmodul **300** wählt die Emphasewörter aus dem durch die Emphasewortauswahlmodul **300** eingegebenen Text aus. Der Informationstypbestimmungsmodul **500** bestimmt den Informationstyp des durch die Domänen-DB **510** eingegebenen Texts und erzeugt die Satzmusterinformation.

[0155] Der Prozess zur Bestimmung des Informationstyps unter Verwendung des Eingabetexts wird ausführlich beschrieben. Die Wörter des Eingabetexts werden gemäß der entsprechenden tatsächlichen Semantikeinheiten unter Verwendung der Morphem- und Semantikstrukturinformation, die vom TTS **100** erhalten ist und der Semantikeinheit-DB der Domänen-DB **510** unterteilt. Das Ergebnis wird wie folgt ausgedrückt:

"/Today,/the Nasdaq composite index/closed/down/0,57/to/1,760.54/ and/the Dow Jones industrial average/finishedup/31.39/to/9397.51./"

[0156] Der Eingabetext wird ausgehend von den vorliegenden Semantikeinheiten unterteilt und die repräsentative Bedeutung wird dann aus dem Eingabetext bestimmt, der ausgehend von den Semantikeinheiten durch Bezugnahme zur Domänen-DB **510** geteilt ist, so dass die bestimmte repräsentative Bedeutung an die Semantikeinheiten angehängt wird. Das Ergebnis der angehängten repräsentativen Bedeutung wird wie folgt ausgedrückt:

"/DATE/INDEX/closed/down/VALUE/to/VALUE/and/INDEX/finished/up/ VALUE/to/VALUE/"

[0157] Dann werden die repräsentativen Wörter der entsprechenden Semantikeinheiten des Eingabetexts ausgewählt, und das Ergebnis der ausgewählten repräsentativen Wörter kann wie folgt ausgedrückt werden:

"/Today/Nasdaq/close/down/0.57/to/1,760.54/and/Dow/finish/up/31.39/to /9397.51./"

[0158] Dann wird die Grammatikregel extrahiert, die auf die Syntax- und Semantikstruktur des Textes von der Domänen-DB **510** angewendet wird und nur der Teil, der der vorliegenden Information im Eingabetext entspricht, wird durch Anwenden der extrahierten Grammatikregel auf den Eingabetext angezeigt, der gemäß der entsprechenden Semantikeinheiten geteilt ist.

[0159] Das heißt, wenn die Syntaxstruktur des Eingabetexts der folgenden Grammatikregel entspricht, die im Informationstyp des aktuellen Status des Börsenmarkts entspricht, wird der Informationstyp des Eingabetexts als der vorliegende Status des Börsenmarkts bestimmt.

– NP{DATE}, NP{INDEX}VP{close PP{*}PP{to VALUE}} → DATE INDEX VALUE
 – NP{INDEX}, VP{finish PP{*}PP{to VALUE}} → INDEX VALUE

[0160] Wenn der Eingabetext mit der extrahierten Grammatikregel angewendet wird, wird der Text wie folgt ausgedrückt:

"INFO[Today/DATE], INFO[the Nasdaq composite index/INDEX] closed down 0.57 to INFO[1,760.54/VALUE] and INFO[the Dow Jones industrial average/INDEX] finished up 31.39 to INFO[9397.51/VALUE]."

[0161] Als Folge davon werden die repräsentativen Wörter (d. h. Today/DATE, Nasdaq/INDEX, 1,760.54/VALUE, DOW/INDEX, 9397.51/VALUE) ausgewählt, die als die tatsächliche Information angezeigt werden sollen. Dann wird ein INDEX-WERT-Typ als Satzmusterinformation zum Anzeigen der repräsentativen Wörter in der am besten geeigneten Weise für den bestimmten Informationstyp erzeugt.

[0162] Wenn die Satzmusterinformation durch den obigen Prozess erzeugt ist, liegt die Satzmusterinformation vor, die auf die vom Emphasewortauswahlmodul **300** ausgewählten Emphasewörter angewendet wird, als Ergebnis der Bestimmung durch die Satzmusterinformationsanpassungseinheit **710** des Strukturierungsmoduls **700**, ob die Satzmusterinformation vorhanden ist. Auf diese Weise wird bestimmt, ob die ausgewählten Emphasewörter auf die Satzmusterinformation angewendet werden können, die vom Informationstypbestimmungsmodul **500** erzeugt ist.

[0163] Wenn die vom Emphasewortauswahlmodul **300** ausgewählten Emphasewörter in den Wörtern enthalten sind, die vom Informationstypbestimmungsmodul **500** als die repräsentativen Wörter ausgewählt sind, die als die tatsächliche Information angezeigt werden sollen, veranlasst die Satzmusteranpassungseinheit **710**, dass die Emphasewörter an die erzeugte Satzmusterinformation angehängt werden.

[0164] Wenn jedoch die ausgewählten Emphasewörter nicht in den Wörtern enthalten sind, die als repräsentative Wörter im Informationstypbestimmungsmodul **500** ausgewählt sind, werden die Emphasewörter gemäß

der Syntaxstruktur des bestimmten Informationstyps umgeordnet.

[0165] Wenn die Emphasewörter an die Satzmusterinformation angehängt sind oder gemäß der Syntaxstruktur in der obigen Weise umgeordnet sind, extrahiert die Informationsstrukturierungseinheit **750** die Metainformation zum Layout der Emphasewörter gemäß dem Informationstyp von der Meta-DB **730** und veranlasst, dass die Emphasewörter an die extrahierte Metainformation angehängt werden.

[0166] Im Prozess zum Veranlassen, dass die Emphasewörter an die Metainformation angehängt werden, werden die entsprechenden synthetisierten Laute, die jedem der Emphasewörter zugeordnet sind, mit der Taktinformation zusammen gesetzt.

[0167] Wenn die Information in der Weise ausgedrückt ist, dass das DATE (Datum) der TITLE (Titel) wird und der INDEX und der VALUE (Wert) in Form einer Tabellenstruktur entsprechend der Positionen im Informationstyp, der dem Börsenmarkt angehört, angegeben werden, wird das als Tabellenform dargestellte Layoutformat aus der Meta-DB **730** extrahiert. Die Emphasewörter und die Taktinformation werden in das extrahierte Layout wie folgt eingegeben:

```
<TITLE SYNC = "510"> Today </TITLE>
<INDEX VALUE ITEM = "2">
<INDEX SYNC = "1351"> Nasdaq </INDEX>
<VALUE SYNC = "INHERIT">1,760.54</VALUE>
.
.
</INDEXVALUE>
```

[0168] Als Folge davon werden, wie in [Fig. 9c](#) gezeigt, die ausgewählten Emphasewörter zusammen mit den entsprechenden synthetisierten Lauten in der Weise angezeigt, dass der VALUE entsprechend den Positionen des Börsenpreisindex zusammen mit dem INDEX durch einen Anhang 'INHERIT' gezeigt wird.

[0169] Gemäß der vorliegenden Erfindung kann der Benutzer die Wörter sichtbar nachprüfen, die für den Benutzer schwierig zu erkennen sind. Auf diese Weise können Beschränkungen bei Zeit und Erkennung, die der Sprache innewohnen, reduziert werden.

[0170] Darüber hinaus kann der Benutzer den Inhalt der in Form synthetisierter Laute bereitgestellten Information durch strukturell angezeigte zusätzliche Information intuitiv besser verstehen. Daraus ergibt sich ein Vorteil, indem die Kapazität und Zuverlässigkeit der Informationsvermittlung der TTS verbessert werden kann.

[0171] Darüber hinaus kann die Betriebseffizienz des Systems zur Umsetzung von Text in Sprache maximiert werden.

[0172] Obwohl die vorliegende Erfindung in Verbindung mit den in den begleitenden Zeichnungen gezeigten Ausführungsformen beschrieben wurde, dient sie nur der Erläuterung. Daher versteht es sich für die Fachleute eindeutig, dass verschiedene Modifikationen und andere Äquivalente daran vorgenommen werden können. Deshalb sollte der Rahmen der vorliegenden Erfindung nur durch die beigefügten Ansprüche definiert sein.

Patentansprüche

1. System zur Umsetzung von Text in Sprache umfassend:
 einen Sprachsynthesemodul zum Analysieren von Textdaten nach Morphemen und einer syntaktischen Struktur, Synthetisieren der Textdaten in Sprache unter Verwendung der erhaltenen Sprachsyntheseanalysedaten und Ausgeben synthetisierter Laute;
 einen Emphasewortwahlmodul zum Auswählen von Wörtern, die zu spezifischen Teilen der Sprache gehören als Emphasewörter aus den Textdaten unter Verwendung der aus dem Sprachsynthesemodul erhaltenen Sprachsyntheseanalysedaten; und
 einen Anzeigemodul zum Anzeigen der ausgewählten Emphasewörter in Synchronisation mit den synthetisierten Lauten.

2. System zur Umsetzung von Text in Sprache nach Anspruch 1, ferner umfassend: einen Informationstypbestimmungsmodul zum Bestimmen des Informationstyps der Textdaten unter Verwendung der vom Sprachsynthesemodul erhaltenen Sprachsyntheseanalysedaten und Erzeugen von Satzmusterinformation; und
worin der Anzeigemodul ferner zum Umordnen der ausgewählten Emphasewörter nach der erzeugten Satzmusterinformation vor Anzeigen der umgeordneten Emphasewörter in Synchronisation mit den synthetisierten Lauten vorgesehen ist.

3. System zur Umsetzung von Text in Sprache nach Anspruch 1 oder 2, ferner umfassend einen Strukturierungsmodul zum Strukturieren der ausgewählten Emphasewörter nach einem bestimmten Layoutformat.

4. System zur Umsetzung von Text in Sprache nach Anspruch 3, worin der Strukturierungsmodul umfasst: eine Meta-DB, in der Layouts zum strukturellen Anzeigen der Emphasewörter, die nach dem Informationstyp ausgewählt sind, und zusätzlich angezeigter Inhalt als Metainformation gespeichert wird; eine Satzmusterinformationsanpassungseinheit zum Umordnen der Emphasewörter, die vom Emphasewortwahlmodul ausgewählt sind, nach der Satzmusterinformation; und eine Informationsstrukturierungseinheit zum Extrahieren von Metainformation entsprechend dem bestimmten Informationstyp aus der Meta-DB und Anwenden der umgeordneten Emphasewörter auf die extrahierte Metainformation.

5. System zur Umsetzung von Text in Sprache nach einem der Ansprüche 1 bis 4, worin die Emphasewörter Wörter beinhalten, von denen erwartet wird, dass sie Verzerrung der synthetisierten Laute bei den Wörtern in den Textdaten aufweisen, unter Verwendung der Sprachsyntheseanalysedaten, die vom Sprachsynthesemodul erhalten sind.

6. System zur Umsetzung von Text in Sprache nach Anspruch 5, worin die Wörter, von denen erwartet wird, dass sie Verzerrung der synthetisierten Laute aufweisen, Wörter sind, bei denen Übereinstimmungsraten geringer sind als ein bestimmter Schwellenwert, wobei jede der Übereinstimmungsraten auf Basis einer Differenz zwischen geschätzter Ausgabe und einem Istwert des synthetisierten Lauts jedes Sprachsegments jedes Worts bestimmt wird.

7. System zur Umsetzung von Text in Sprache nach Anspruch 6, worin die Differenz zwischen geschätzter Ausgabe und Istwert nach der folgenden Gleichung berechnet wird:

$$\Sigma Q \text{ (Größe der (Eingabe), } |\text{Schätzwert} - \text{Istwert}|, C)/N,$$

wo C ein Übereinstimmungswert (Konnektivität) ist und N ein normalisierter Wert (Normalisierung) ist.

8. System zur Umsetzung von Text in Sprache nach einem der Ansprüche 1 bis 4, worin die Emphasewörter ausgewählt sind aus Wörtern, deren Emphasfrequenzen geringer sind als ein bestimmter Schwellenwert unter Verwendung von Information zu Emphasfrequenzen für die entsprechenden Wörter in den Textdaten erhalten vom Sprachsynthesemodul.

9. Verfahren zur Umsetzung von Text in Sprache umfassend die Schritte:
einen Sprachsyntheseschritt zum Analysieren von Textdaten nach Morphemen und einer syntaktischen Struktur, Synthetisieren der Textdaten in Sprache unter Verwendung erhaltener Sprachsyntheseanalysedaten und Ausgeben synthetisierter Laute;
einen Emphasewortauswahlschritt zum Auswählen von Wörtern, die zu spezifischen Teilen von Sprache gehören als Emphasewörter aus den Textdaten unter Verwendung der Sprachsyntheseanalysedaten; und
einen Anzeigeschritt zum Anzeigen der ausgewählten Emphasewörter in Synchronisation mit den synthetisierten Lauten.

10. Verfahren zur Umsetzung von Text in Sprache nach Anspruch 9, wobei das Verfahren nach dem Emphasewortauswahlschritt und vor dem Anzeigeschritt ferner umfasst:
einen Satzmusterinformationserzeugungsschritt zum Bestimmen des Informationstyps der Textdaten unter Verwendung der vom Sprachsyntheseschritt erhaltenen Sprachsyntheseanalysedaten und Erzeugen von Satzmusterinformation; und
worin der Anzeigeschritt ferner zum Umordnen ausgewählter Emphasewörter nach der erzeugten Satzmusterinformation vorgesehen ist vor einem Anzeigen der umgeordneten Emphasewörter in Synchronisation mit den synthetisierten Lauten.

11. Verfahren zur Umsetzung von Text in Sprache nach Anspruch 9 oder 10, ferner umfassend einen Strukturierungsschritt zum Strukturieren der ausgewählten Emphasewörter nach einem bestimmten Layoutformat.

12. Verfahren zur Umsetzung von Text in Sprache nach Anspruch 11, worin der Strukturierungsschritt die Schritte umfasst:

Bestimmen, ob die ausgewählten Emphasewörter auf den Informationstyp der erzeugten Satzmusterinformation anwendbar sind;

Veranlassen, dass die Emphasewörter in der Satzmusterinformation markiert werden nach einem Ergebnis des Bestimmungsschritts oder Umordnen der Emphasewörter nach dem bestimmten Informationstyp; und Strukturieren der umgeordneten Emphasewörter nach der Metainformation entsprechend dem aus der Meta-DB extrahierten Informationstyp.

13. Verfahren zur Umsetzung von Text in Sprache nach Anspruch 12, worin Layouts zur strukturellen Anzeige der Emphasewörter ausgewählt nach dem Informationstyp und zusätzlich angezeigte Inhalte als Metainformation in der Meta-DB gespeichert werden.

14. Verfahren zur Umsetzung von Text in Sprache nach einem der Ansprüche 9 bis 13, worin der Emphasewortauswahlschritt ferner den Schritt umfasst zum Auswählen von Wörtern, von denen erwartet wird, dass sie Verzerrung der synthetisierten Laute aufweisen, aus Wörtern in den Textdaten unter Verwendung der im Sprachsyntheseschritt erhaltenen Sprachsyntheseanalysedaten.

15. Verfahren zur Umsetzung von Text in Sprache nach Anspruch 14, worin die Wörter, von denen erwartet wird, dass sie Verzerrung der synthetisierten Laute aufweisen, Wörter sind, deren Übereinstimmungsraten geringer sind als ein bestimmter Schwellenwert, wobei jede der Übereinstimmungsraten auf Basis einer Differenz zwischen geschätzter Ausgabe und einem Istwert des synthetisierten Lauts jedes Sprachsegments jedes Worts bestimmt wird.

16. Verfahren zur Umsetzung von Text in Sprache nach einem der Ansprüche 9 bis 13, worin im Emphasewortauswahlschritt die Emphasewörter aus Wörtern ausgewählt werden, deren Emphasefrequenzen geringer sind als ein bestimmter Schwellenwert unter Verwendung von Information zu Emphasefrequenzen für entsprechende Wörter in den Textdaten erhalten vom Sprachsyntheseschritt.

17. Verfahren zur Umsetzung von Text in Sprache nach Anspruch 10, worin der Satzmusterinformationserzeugungsschritt die Schritte umfasst:

Unterteilen der Textdaten in semantische Einheiten unter Bezugnahme auf eine Domänen-DB und die im Sprachsyntheseschritt erhaltenen Sprachsyntheseanalysedaten;

Bestimmen repräsentativer Bedeutungen der unterteilten semantischen Einheiten, Markieren der repräsentativen Bedeutungen der semantischen Einheiten und Auswählen repräsentativer Wörter aus den entsprechenden semantischen Einheiten;

Extrahieren einer Grammatikregel, die für ein syntaktisches Strukturformat des Texts aus der Domänen-DB geeignet ist, und Bestimmen von Istinformation zum Anwenden der extrahierten Grammatikregel auf die Textdaten; und

Bestimmen des Informationstyps der Textdaten durch die bestimmte Istinformation und Erzeugen der Satzmusterinformation.

18. Verfahren zur Umsetzung von Text in Sprache nach Anspruch 17, worin Information zu einer syntaktischen Struktur, einer Grammatikregel, Terminologien und Phrasen verschiedener Bereiche unterteilt nach dem Informationstyp als Domäneninformation in der Domänen-DB gespeichert werden.

Es folgen 10 Blatt Zeichnungen

FIG. 1

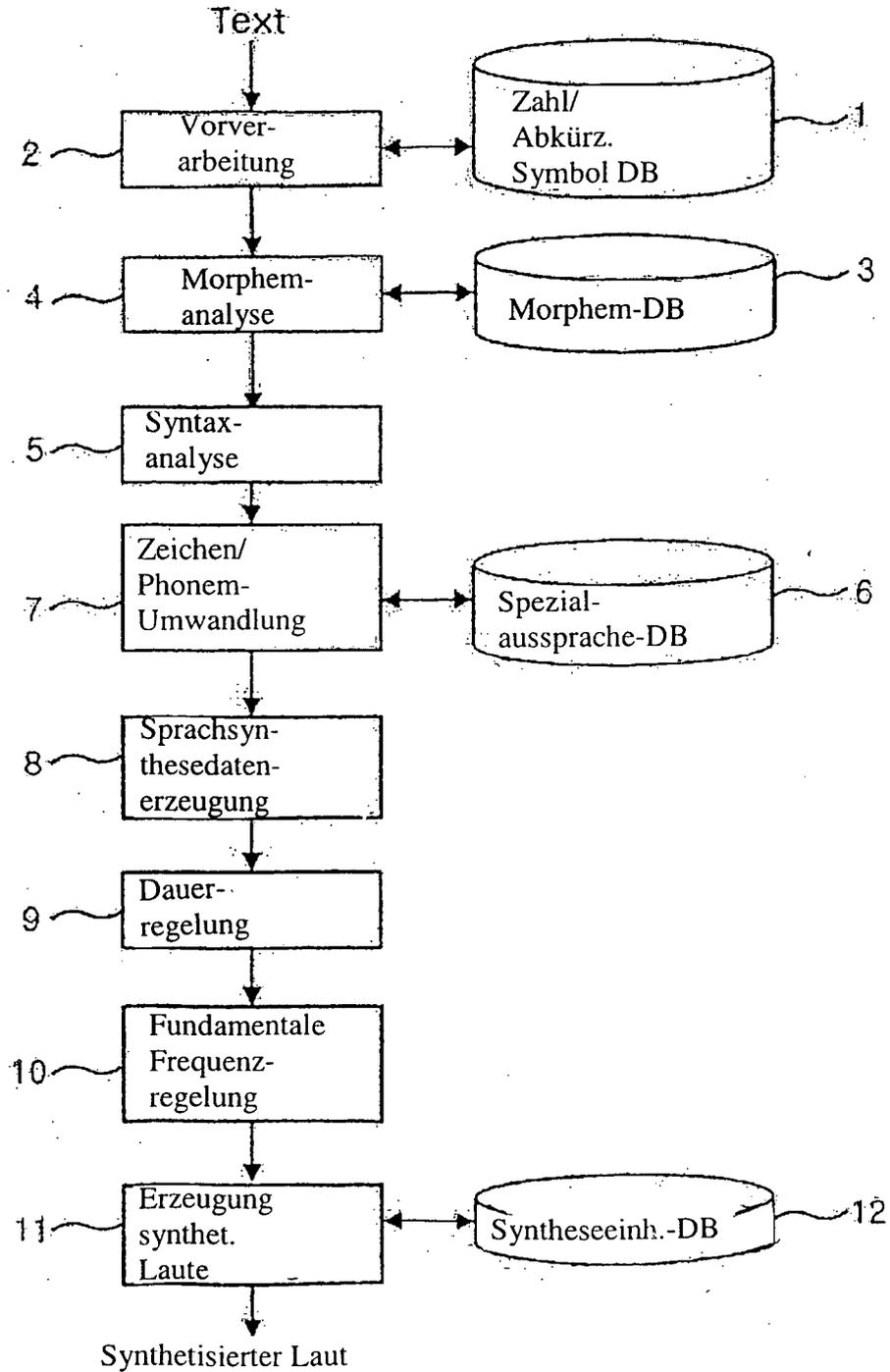


FIG. 2

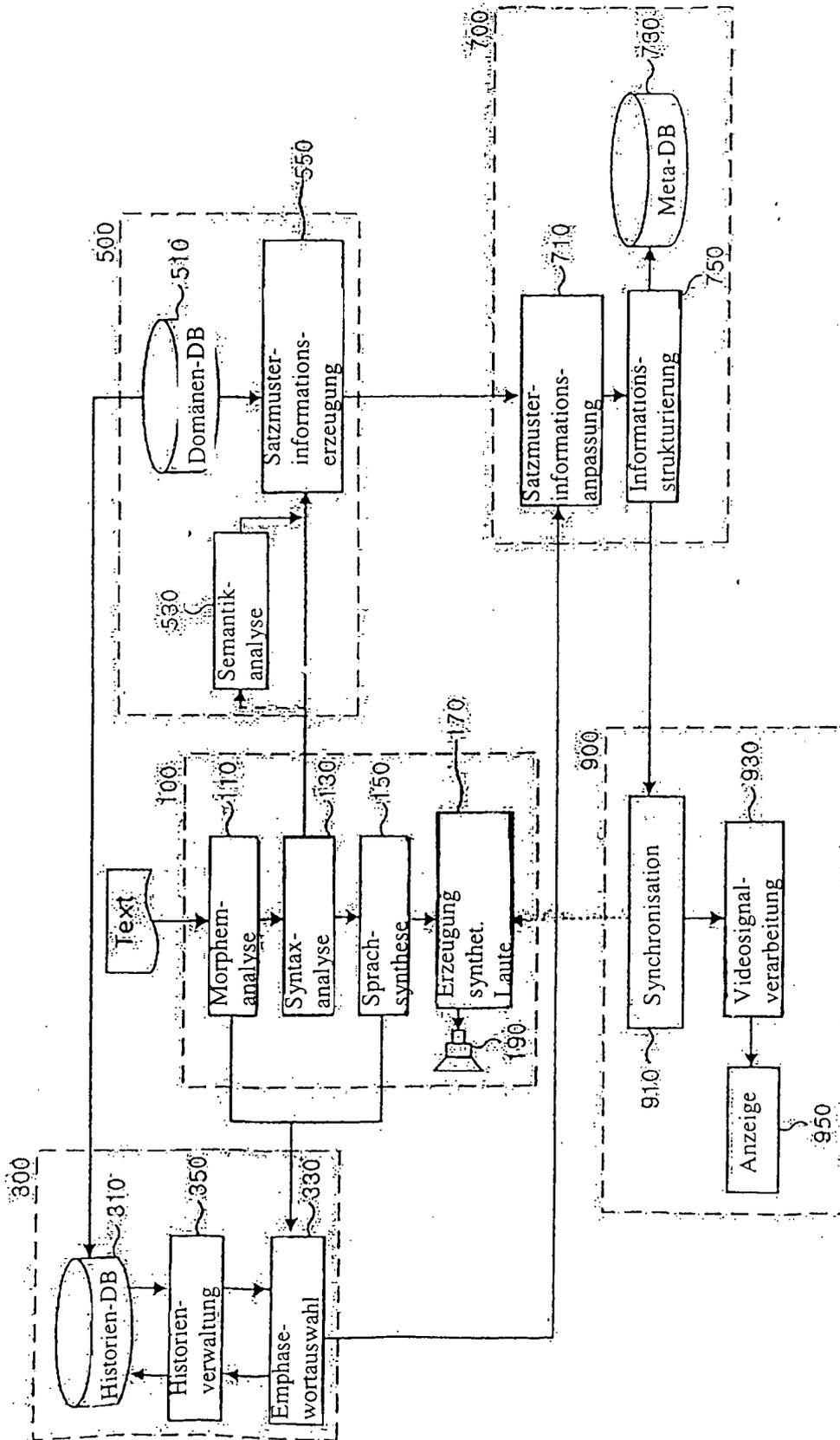


FIG. 3

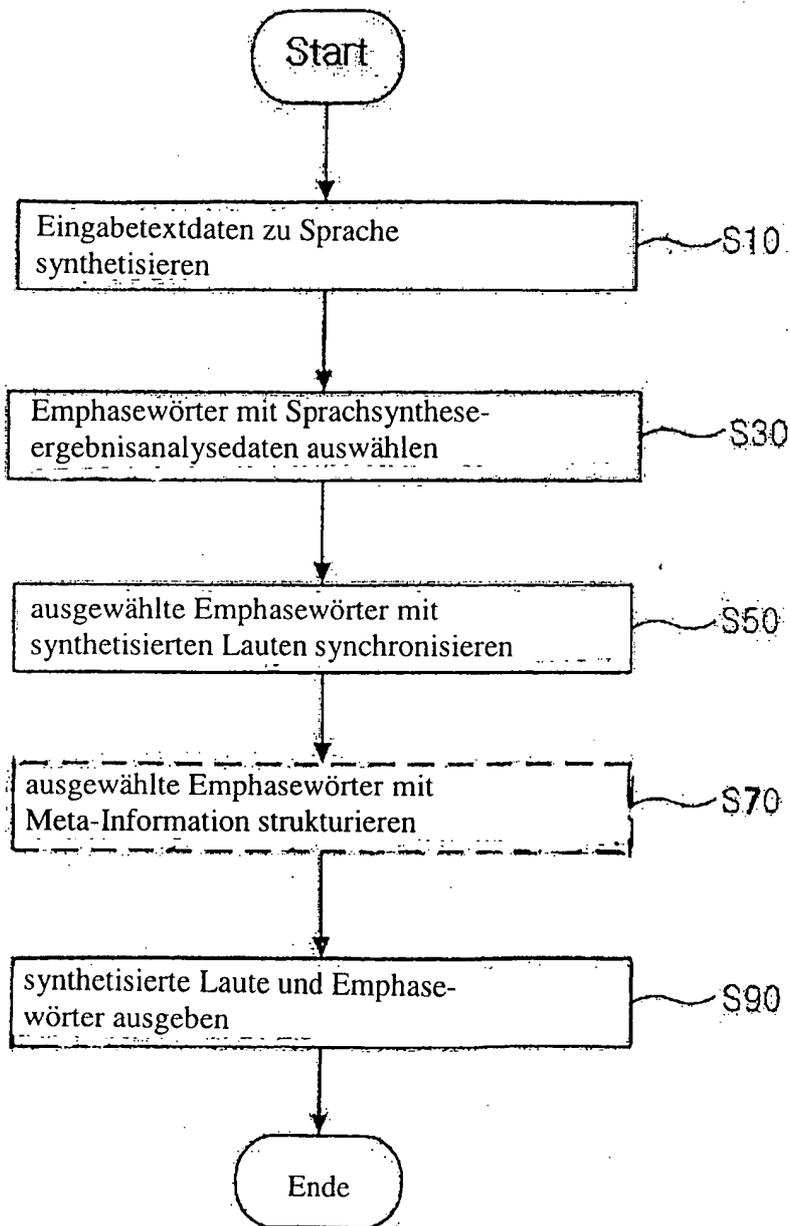


FIG. 4

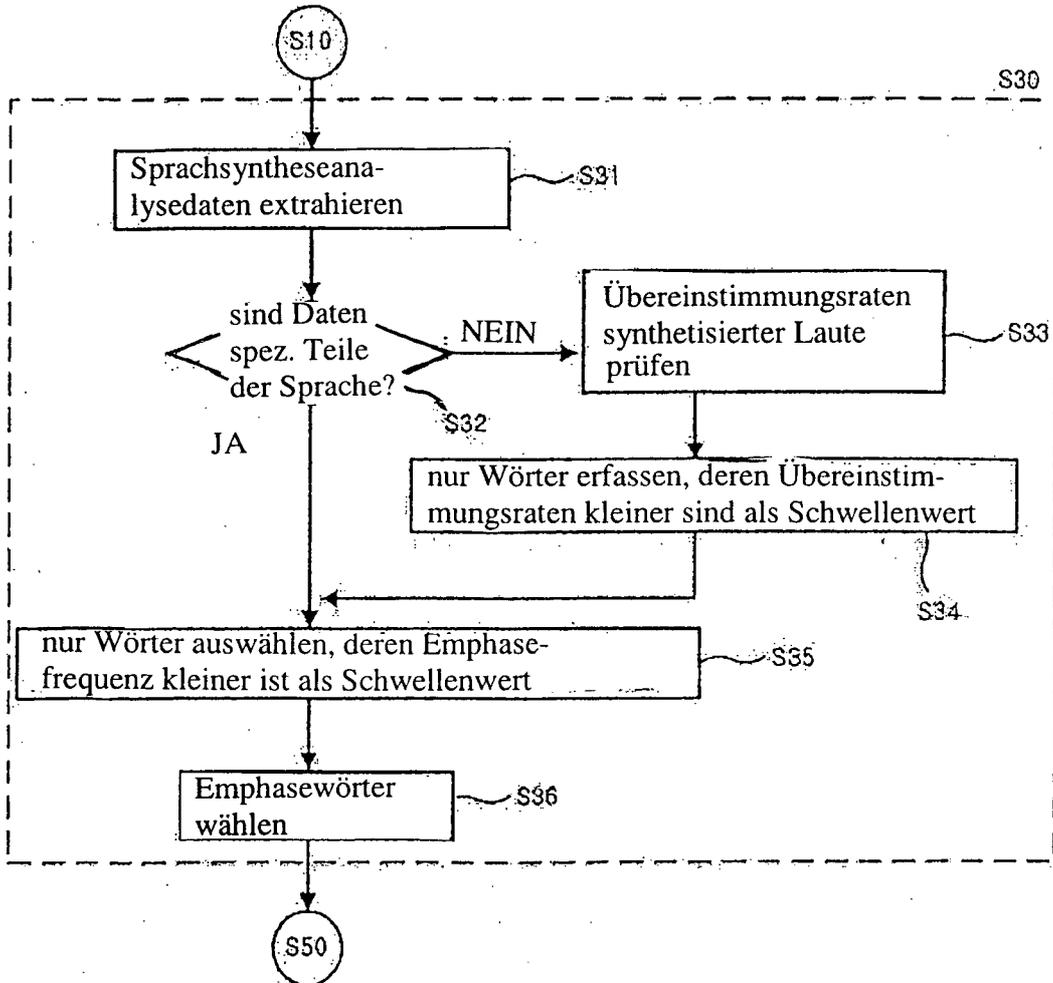


FIG. 5

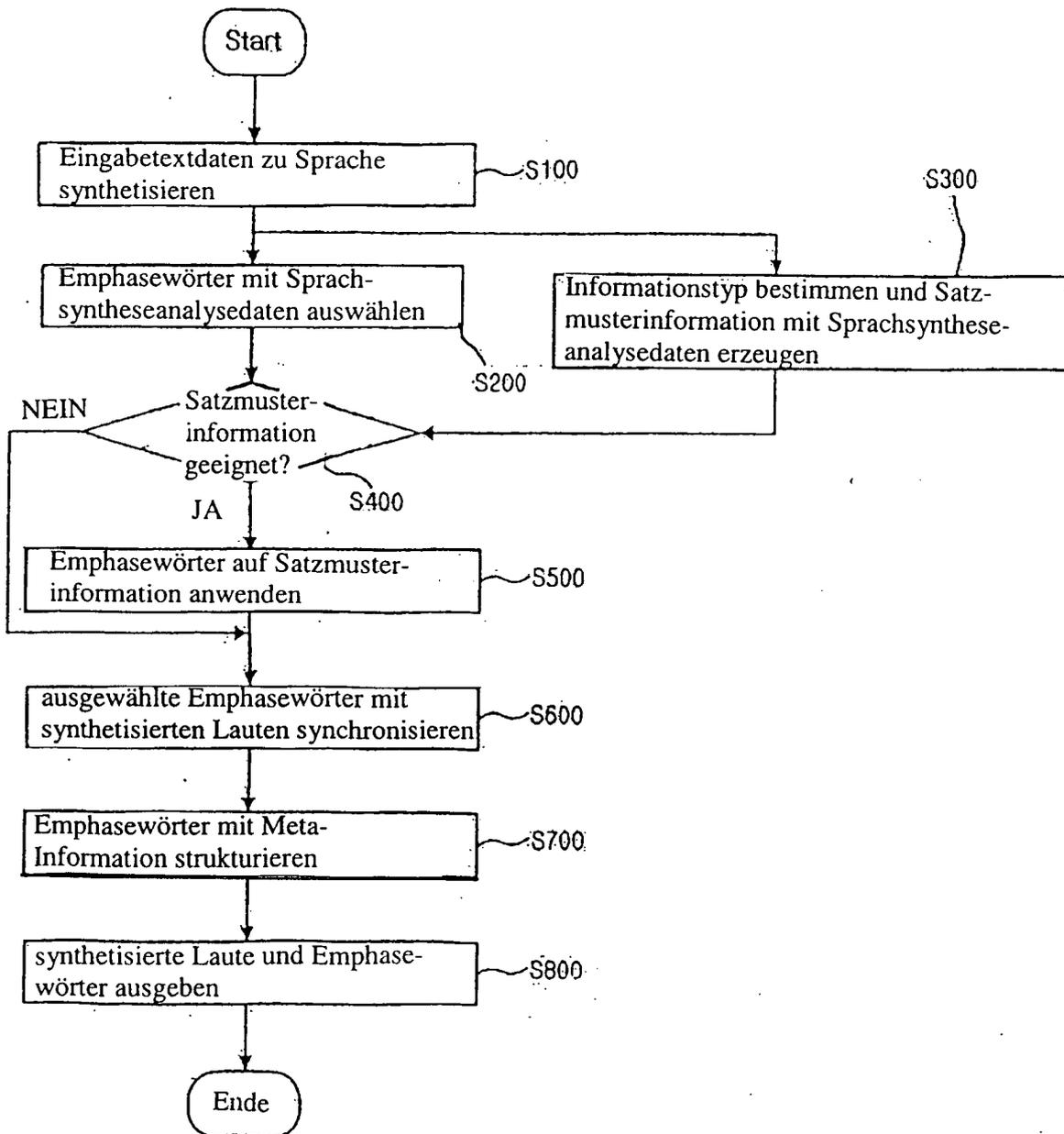


FIG. 6

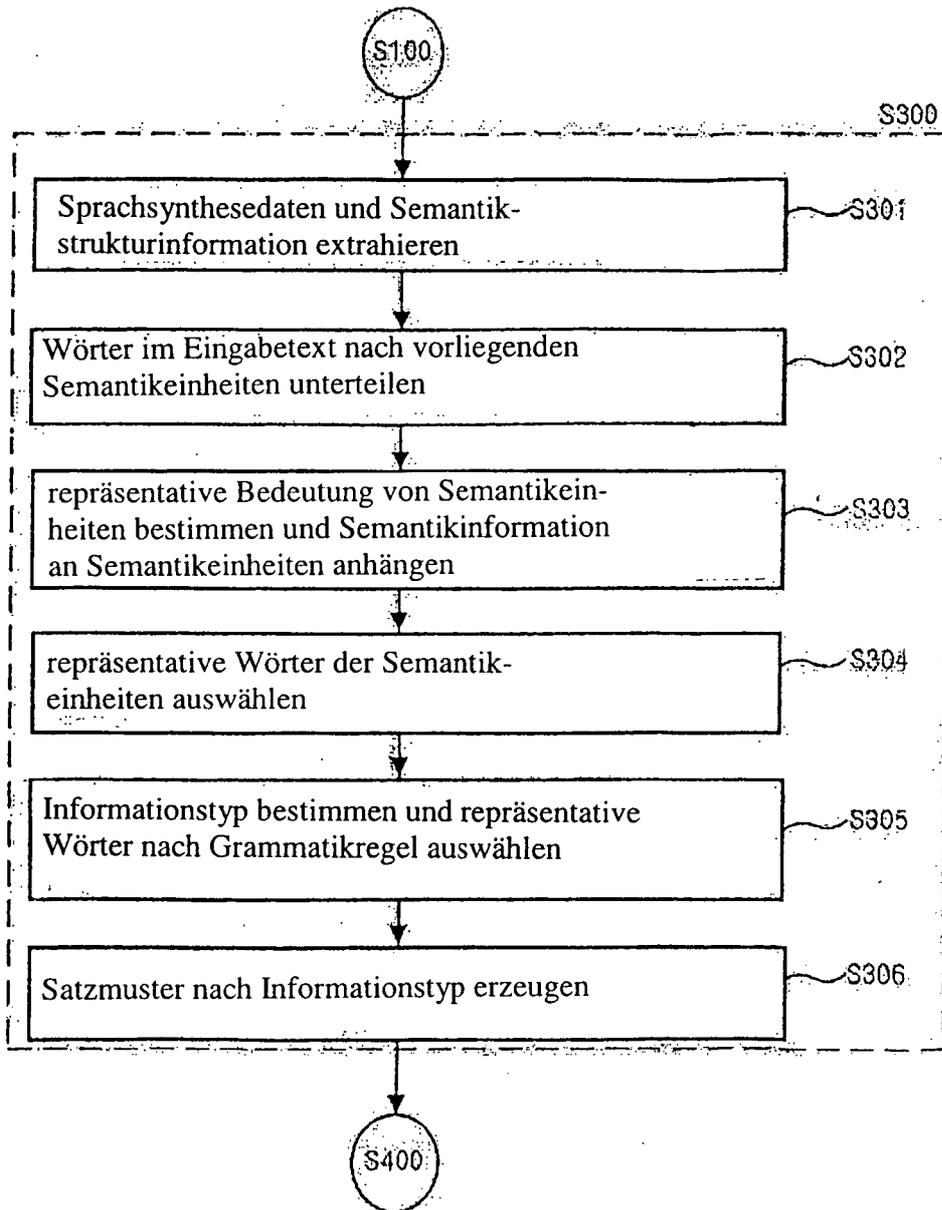


FIG. 7

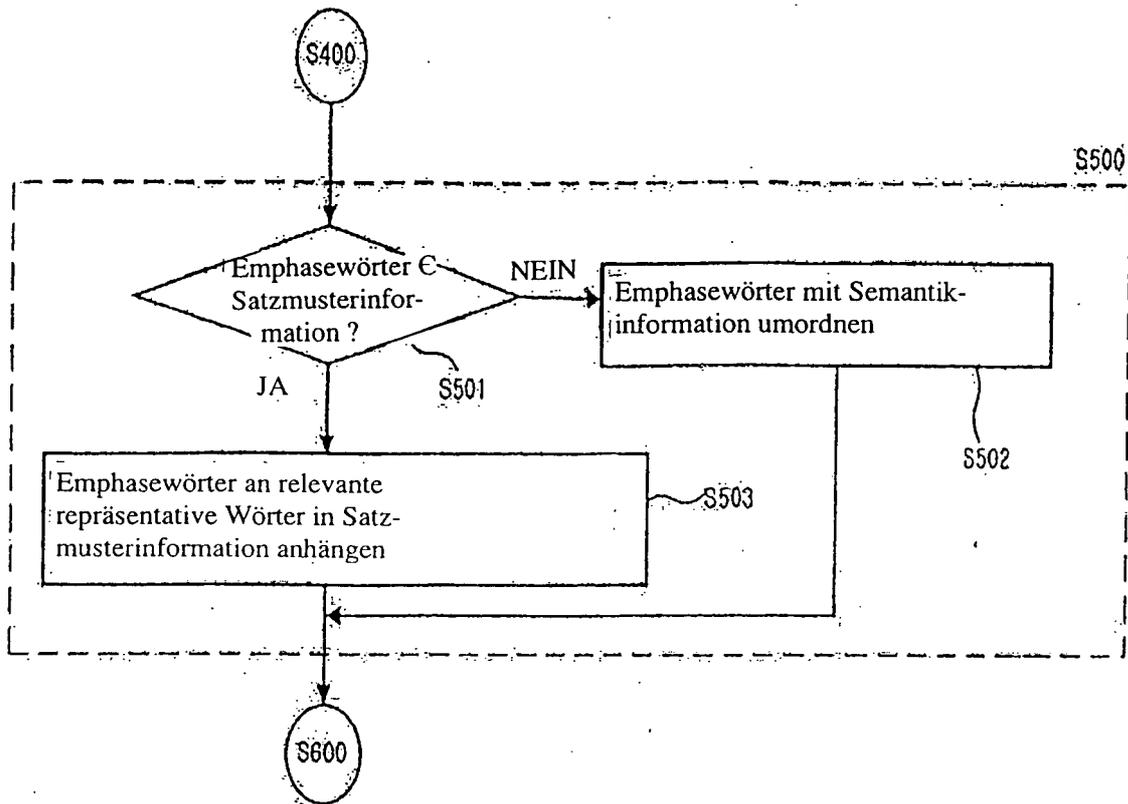


FIG. 8

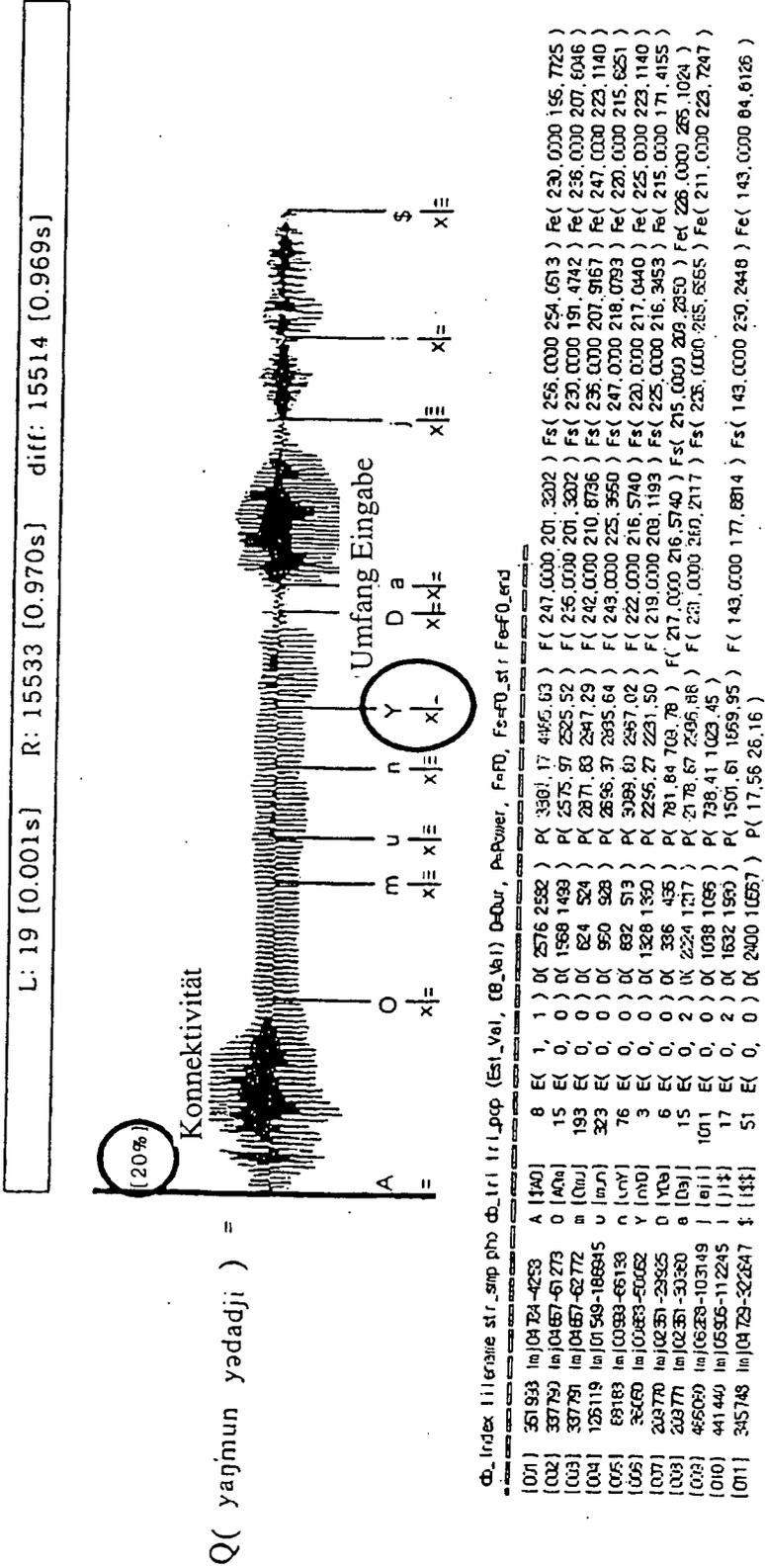


FIG. 9a

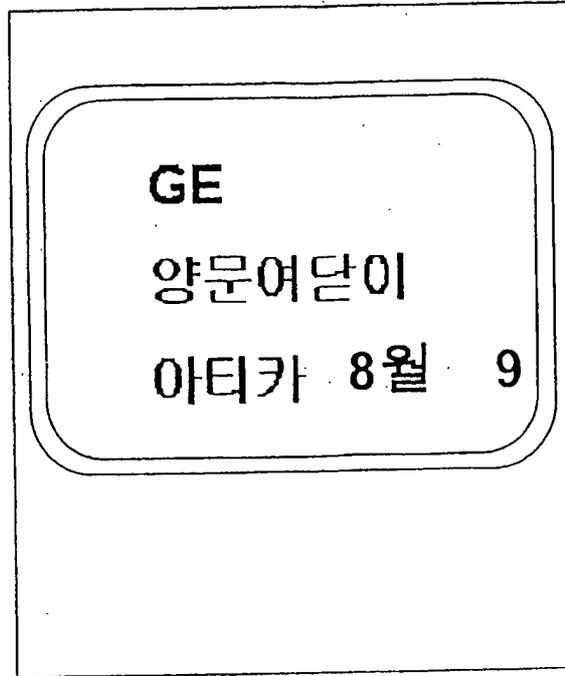


FIG. 9b

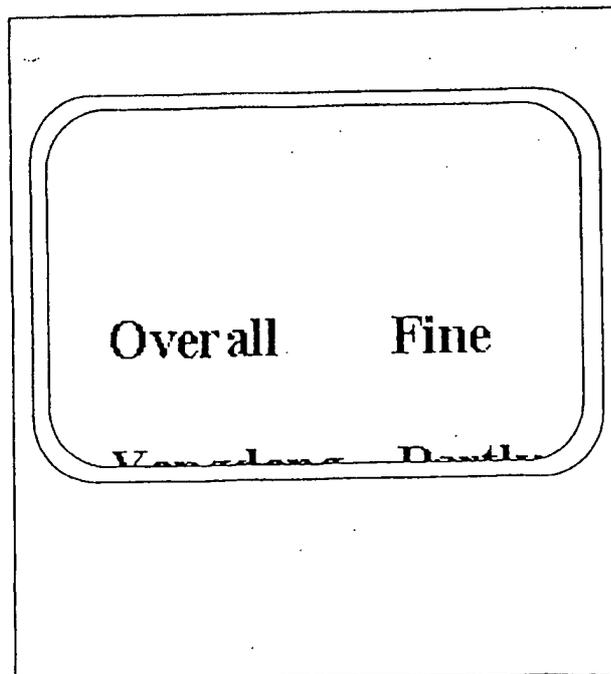
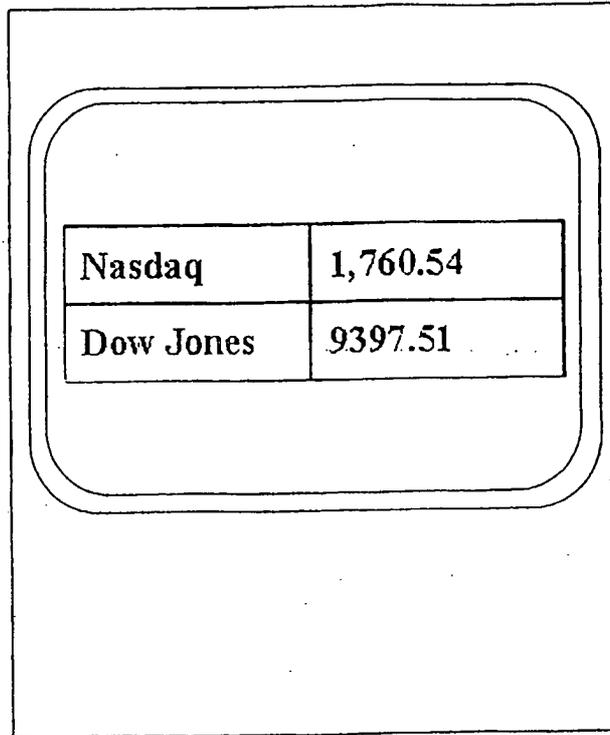


FIG. 9c



Nasdaq	1,760.54
Dow Jones	9397.51