(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Veröffentlichungsnummer: **0 058 130**
A2

(12) EUROPÄISCHE PATENTANMELDUNG

(21) Anmeldenummer: 82730011.2

(22) Anmeldetag: 11.02.82

(51) Int. Cl.³: G 10 L 1/08

(30) Priorität: 11.02.81 DE 3105518

(43) Veröffentlichungstag der Anmeldung: 18.08.82 Patentblatt 82/33

(84) Benannte Vertragsstaaten: AT CH DE FR GB LI

(71) Anmelder: Heinrich-Hertz-Institut für Nachrichtentechnik Berlin GmbH
Einsteinufer 37
D-1000 Berlin 10(DE)

(72) Erfinder: Grossmann, Eberhard
Wüllenweberweg 3
D-1000 Berlin 45(DE)

(74) Vertreter: Wolff, Konrad
Heinrich-Hertz-Institut für Nachrichtentechnik Berlin GmbH Einsteinufer 37
D-1000 Berlin 10(DE)

(54) Verfahren zur Synthese von Sprache mit unbegrenztem Wortschatz und Schaltungsanordnung zur Durchführung des Verfahrens.

(57) Unter Sprachsynthese ist die Umwandlung von Texten, die als Symbolfolgen eingegeben werden, in Folgen äquivalenter akustischer Signale zu verstehen. Bei der Erfindung geschieht dies für einen unbegrenzten Wortschatz und im Hinblick auf den erforderlichen Aufwand sowie auf die erzielbare Verständlichkeit und Natürlichkeit im Zeitbereich mit knapp 100 Lautelementen, die etwa 22 kByte Speicherplatz benötigen. Es sind etwa 40 Elemente für Einzellaute und etwa 50 Elemente für Übergangslaute vorgesehen. Elemente stimmhafter Einzel- und Übergangslaute weisen vorgegebene, spezielle Abtastwerte auf, die zur Tonhöhenänderung ausgelassen bzw. mindestens einmal verwendet werden können. Dabei bleibt der Lautcharakter erhalten.

Als Eingangsbefehle benötigt das Sprachsynthesesystem an sich Lautschriftzeichen.

Um die Anwendung zu erleichtern und sie auch ungeübten Benutzern zu eröffnen, ist ein Transkriptionssystem vorgesehen, das mit üblichen Schriftzeichen eingegebene Texte in einem der Synthese unmittelbar vorausgehenden Schritt selbsttätig in die erforderlichen Lautschriftzeichenfolgen umwandelt.

EP 0 058 130 A2

HEINRICH-HERTZ-INSTITUT FÜR NACHRICHTENTECHNIK BERLIN GMBH

01/0281 EP

Verfahren zur Synthese von Sprache mit unbegrenztem Wortschatz und Schaltungsanordnung zur Durchführung des Verfahrens

Die Erfindung bezieht sich auf ein Verfahren zur Synthese von Sprache mit unbegrenztem Wortschatz im Zeitbereich aus Lautelementen, die aus natürlichen Sprachproben gewonnen und in digitaler Form, redundanzarm kodiert, gespeichert und außerdem im Hinblick auf den erforderlichen Speicherplatzbedarf in der Länge jeweils auf den signifikanten Bereich des betreffenden lauttypischen Zeitsignals und in der Anzahl unter Ausnutzung sich gegenseitig ineinander überführbarer verwandter Laute reduziert sind, wobei zur Sprachsynthese diese Lautelemente aufgrund von Eingangsbefehlen und von vorgegebenen Verknüpfungsregeln in der erforderlichen Gestalt, Anzahl und Reihenfolge zu digitalen Signalfolgen verkettet werden, aus denen mittels Digital-Analog-Wandlung und steuerbarer Verstärkung als Sprache wahrnehmbare Schallwellen erzeugt werden, sowie auf eine Schaltungsanordnung zur Durchführung des Verfahrens.

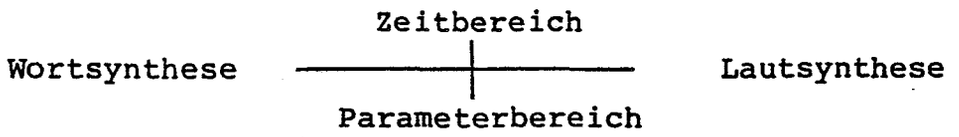
Unter Sprachsynthese ist die Umwandlung eines als Symbolfolge vorliegenden Textes in das äquivalente akustische Signal mittels einer technischen Apparatur zu verstehen. Dabei ist es von grundlegender Bedeutung, daß zwischen der Eingabe der Symbolfolge in die Apparatur und der Ausgabe des äquivalenten akustischen Signals alle Abläufe unmittelbar, ohne Zwischenschaltung menschlicher Verstandeskräfte stattfinden. Die genau bestimmten technischen Einzelmaßnahmen folgen dabei einem planmäßigen Einsatz berechen- und beherrschbarer Naturkräfte.

Die Bewertungskriterien für synthetische Sprache sind die Verständlichkeit und die Natürlichkeit. Die Maßstäbe dafür sind, wenn auch z.B. bei der Verständlichkeit nach objektiven Gesichtspunkten feststellbar, subjektiver Natur. Dennoch gibt es Sachverhalte, die für die Beurteilung sofort von jedermann herangezogen werden. Dabei handelt es sich um den Verlauf der Grundtonhöhe (Pitchfrequenz), den Sprechrhythmus und um den Intensitätsverlauf. Beim Signalverlauf natürlicher Sprache gehen die Einzellaute ineinander über. Sie werden durch mehrere Lautbildungsfrequenzen (Formanten) charakterisiert. Diese Lautbildungsfrequenzen sind unabhängig von der Grundtonhöhe, d.h. unabhängig von der Sprechhöhe. Diese Sachverhalte wirken sich mehr oder weniger sowohl auf die Verständlichkeit als auch auf die Natürlichkeit aus. Während die Verständlichkeit bei bekannten Sprachsynthesesystemen bisher notgedrungen im Vordergrund stand, zielen die Bestrebungen neuerdings, nachdem eine ausreichende Verständlichkeit erreicht wurde, mehr und mehr auf Verbesserungen hinsichtlich der Natürlichkeit ab. Geringe Schwierigkeiten bestehen bei der Dynamik. Die relative Lautstärke läßt sich mit steuerbaren Verstärkern variieren. Auch die Lautdauer, und damit der Sprechrhythmus, läßt sich durch dynamische Steuerung der Wiederholanzahl der Einzellautelemente mit verhältnismäßig einfachen Mitteln verändern. Problematisch hingegen ist die Beherrschung der Melodik, da die Länge der Sprachgrundfrequenzperioden für die einzelnen Laute fest vorgegeben sind und eine einfache, proportionale Verlängerung oder Verkürzung von Sprachgrundfrequenzperioden eine entsprechende Verschiebung des Formantenfrequenzspektrums bedeutet, d.h. zur völlig unnatürlichen Lauten führt.

Verständlichkeit und Natürlichkeit synthetischer Sprache hängen andererseits auch von der Leistung ab, für das das betreffende System konzipiert ist. Selbstverständlich kann bei einem System mit begrenztem Wortschatz eine hervorragende Qualität der Sprache gewährleistet werden. Komplette

Wörter oder gar längere Phrasen, zudem vielleicht noch von einem geschulten Sprecher dargeboten, können unter Erhaltung der natürlichen Melodik und Rhythmik gespeichert und auf Abruf wiedergegeben werden. Besteht die Zielsetzung für ein Sprachsynthesesystem hingegen darin, einen unbegrenzten Wortschatz zu erzeugen, muß auf kleinere Synthesebausteine, z.B. auf Laute zurückgegriffen werden. Dabei gehen auf jeden Fall Satz- und Wortdynamik sowie die Melodik zunächst verloren und sind bei der Synthese neu zu generieren. In welchem Umfang dies gelingt, ist für die Natürlichkeit synthetischer Sprache von wesentlicher Bedeutung.

Hier nun spielen die technischen Möglichkeiten und die wirtschaftlichen Gesichtspunkte eine ausschlaggebende Rolle. Eine Klassifizierung der Synthesesysteme bzw. deren Unterteilung nach dem Syntheseprinzip



ermöglicht eine erste Abschätzung des erforderlichen Aufwandes für die Realisierung: Eine Wortsynthese, sowohl im Zeitbereich als auch im Parameterbereich, benötigt mit wachsendem Umfang des auszugebenden Vokalubars auch ein wachsendes Speichervolumen. Derartige Systeme sind also mit vernünftigem Aufwand nur für Systeme mit begrenztem Wortschatz geeignet. Auf der Lautsynthese beruhende Systeme ermöglichen die Ausgabe eines unbeschränkten Vokalubars und erfordern unterschiedlichen Aufwand, der in der folgenden Tabelle grob angedeutet ist.

Kriterium	Zeitbereich	Parameterbereich
Algorithmus	+ einfach	- aufwendig
Synthetisier-geschwindigkeit	+ hoch	- mäßig
Speicherplatz-bedarf	- groß	+ gering
Melodievariation	- bisher schwierig	+ leicht

In der technisch-wissenschaftlichen und der Patentliteratur sind die verschiedenartigen Sprachsynthesesysteme in großer Zahl abgehandelt. So ist beispielsweise aus der DE-OS 30 06 339 ein Verfahren und eine Einrichtung zur Sprachsynthese bekannt, wobei zum Zwecke der Miniaturisierung eine Informations-Kompressionstechnik zur Anwendung kommen soll, die bei minimalem Verlust an Sprachverständlichkeit und Natürlichkeit eine Speicherung in einem einzigen integrierten LSI-Schaltungs-Chip möglich werden läßt. Die als Synthesebausteine abgespeicherten Phoneme (Einzellaute) sind bei der Synthese in ihrer aus dem Speicher abgerufenen Gestalt einer Veränderung oder Regulierung in bezug auf eine Anpassung des Tonhöhenintervalls, der Amplituden und der Zeitachse zu unterziehen, um sich der Qualität der natürlichen Sprache wieder anzunähern. Die angewendete Datenkompressionstechnik, die an einem Beispiel näher erläutert ist, führt dazu, daß für ein Wort (Beispiel: "nana") eine Folge weniger (im Beispiel: fünf) Phoneme abzuspeichern ist. Diese, an sich bekannten Tatsachen werden in dieser Vorveröffentlichung detailliert beschrieben. Es ist jedoch kein Hinweis darauf zu entnehmen, ob Möglichkeiten vorgesehen sind, einen unbegrenzten Wortschatz zu synthetisieren sowie Melodik und Rhythmik nach Belieben zu beeinflussen.

Das aus der DE-OS 20 16 572 bekannte Sprachsynthesesystem berücksichtigt insbesondere hinsichtlich der Verständlichkeit die Probleme an den Übergängen zwischen aufeinanderfolgenden Phonemen. Da die Formantfrequenzen - eine Berücksichtigung der drei Hauptformanten ist ausreichend - an den

Übergängen zunehmen, abnehmen oder gleich bleiben können, ergeben sich rein rechnerisch für jedes abzuspeichernde Phonem neun Versionen. Um nicht die Speicherkapazität um praktisch eine weitere Zehnerpotenz erhöhen zu müssen, zielt die Lösung bei diesem bekannten Stand der Technik darauf ab, mit einer gespeicherten Version auszukommen und diese Darstellung den Erfordernissen entsprechend während des Synthesevorgangs zu modifizieren. Außerdem wird lediglich der signifikante Bereich der einzelnen Laute abgespeichert, der z.B. bei einem /s/-Laut nur 10 % der gesamten Lautdauer betragen muß und dementsprechend durch zehnmaliges Wiederholen genau genug und verständlich reproduzierbar ist. Zur Vermeidung von abrupten Übergängen zwischen zwei aufeinanderfolgenden Phonemen sollen die gespeicherten Abschnitte mit einem Schwingungs-Nulldurchgang beginnen. Für stimmhafte Phoneme ist außerdem die Eignung am Übergang zu anderen Phonemen in besonderer Weise - einer subjektiven Prüfung - auszuwählen. Durch diesen Kompromiß lassen sich zwar abrupte Übergänge vermeiden oder zumindest auf einen geringen Umfang reduzieren, wobei jedoch andererseits auf völlig stoßfreie Übergänge verzichtet werden muß.

Dem aus der DE-OS 23 06 816 bekannten Sprachgenerator liegt als Aufgabenstellung bei der Aufbereitung phonetischer Segmente zugrunde, einen umfassenden Tonhöhenperioden-Regelbereich der synthetisierten Laute zu schaffen, der der Verbesserung der Natürlichkeit und der Verständlichkeit zugute kommen soll. Als Lösung wird dazu angegeben, bei stimmhaften Lauten mit definierter Periodizität jeder Tonhöhenlänge Laut-Wellenformen aus natürlicher Sprache herauszugreifen und jeder solchen Wellenform am Endbereich eine Wellenform hinzuzufügen, die durch eine überschlägige Rechnung für die Wellenform des jeweiligen Lauts gewonnen wurde. Laut-Wellenformen von stimmlosen Lauten und die Übergänge zwischen Konsonanten und Vokalen, die eine undefinierte Periodizität aufweisen, sollen in feste Längen unterteilt werden. Die so gewonnenen Laut-Wellenformen stellen dann die Synthesebau-

steine dar. Eine Veränderung der Dauer einer Pitchperiode hat aber nicht nur eine entsprechende Tonhöhenveränderung, sondern - wie bereits oben schon erwähnt und auch nachfolgend noch näher erläutert wird - auch eine Lautverschiebung bzw. eine Verunreinigung zur Folge.

Bei der Erfindung wird von einem Stand der Technik ausgegangen, wie er aus der DE-OS 25 31 006 bekannt und im Oberbegriff des Anspruches 1 berücksichtigt ist. Die danach bei guter Verständlichkeit mögliche Reduktion führte bereits zu einem benötigten Speichervolumen für die Speicherung der Sprachdaten, unkodiert, im Zeitbereich von nur noch ca. 1 Mbit, entsprechend 125 kByte. Ziel der Erfindung ist nun, den Speicherplatzbedarf weiter zu verringern und insbesondere im Hinblick auf die Natürlichkeit der zu synthetisierenden Sprache einfach beherrschbare Maßnahmen zur Wort- und Satz-Melodievariation anzugeben, womit die der Sprachsynthese im Zeitbereich innewohnenden Vorzüge in bezug auf die Verständlichkeit, den Synthesealgorithmus und die Synthetisiergeschwindigkeit erheblich an Bedeutung gegenüber den im Parameterbereich arbeitenden Systemen gewinnen. Gemäß der Erfindung wird dies dadurch erreicht, daß insgesamt ca. 100 Lautelemente vorgesehen sind, nämlich:

- etwa 50 Elemente für Übergangslaute mit je durchschnittlich 240 Abtastwerten für 8 kHz Ausgabefrequenz

und

- etwa 40 Elemente für Einzellaute mit je durchschnittlich 500 Abtastwerten bei stimmlosen und 140 Abtastwerten bei stimmhaften Einzellauten und 8 kHz Ausgabefrequenz,

und daß die Tonhöhe für die Wiedergabe bei den Elementen für die stimmhaften Übergangs- und Einzellaute veränderbar ist, indem solche Abtastwerte, die an diskreten Stellen des Zeitsignals mittels Markierwörtern als geeignet vorgegeben sind, je nach Bedarf aufgrund entsprechender Eingangsbefehle bei der Bildung der digitalen Signalfolgen ausgelassen bzw. mindestens einmal verwendet werden.

Ohne die Bedeutung der angegebenen Einzelheiten bei der Reduzierung der Sprachdaten schmälern zu wollen, werden nachfolgend zunächst die Maßnahmen für die Melodievariation näher erläutert. Wesentlich dafür ist die Tatsache, daß Veränderungen der Melodie von Sprache auf die stimmhaften Anteile entfallen und daß stimmhafte Laute eine große Periodizität aufweisen. Die zu speichernden signifikanten Bereiche benötigen also nur verhältnismäßig wenig wahre Abtastwerte, in der Größenordnung von 80 wahren Abtastwerten je stimmhaften Einzellaut. Innerhalb dieser signifikanten Bereiche, die eine Pitchperiode darstellen und das lauttypische Frequenzgemisch der Formanten enthalten, gibt es mehrere diskrete Stellen, an denen das Formantenfrequenzgemisch im Zeitsignalverlauf kaum oder nur geringfügige Veränderungen zeigt. Die für die Erfindung wesentliche Erkenntnis liegt nun darin, genau aus diesen "unempfindlichen" diskreten Stellen bewußt Veränderungsmöglichkeiten vorzusehen. Das bedeutet, die Pitchperiode kann verändert, verlängert oder verkürzt, und damit die Grundtonhöhe entsprechend abgesenkt oder angehoben werden, wenn Abtastwerte an diesen diskreten Stellen verwendet oder ausgelassen werden, ohne daß sich dadurch der Lautcharakter ändert. Zur Lokalisierung dieser diskreten Stellen, etwa 30 innerhalb eines derartigen signifkanten Bereiches, dienen besondere "Abtastwerte", die Markierwörter, die es erlauben, diese Stellen jederzeit aufzufinden. Die Markierwörter selbst entfallen bei der Verkettung der Elemente zu den digitalen Signalfolgen. Entsprechend dazu lassen 60 Abtastwerte, z.B. die jeweils einem Markierwort benachbarten, je nachdem, ob sie verwendet werden oder nicht, eine praktisch kontinuierliche Variation der Tonhöhe, also sehr viele Melodieverläufe zu. Insbesondere lassen sich dadurch auch die Sprachgrundfrequenzverläufe an den Übergängen zu den folgenden Lauten kontinuierlich gestalten, also Stoßstellen vermeiden.

Hierin liegt auch ein Grund dafür, daß als Synthesebausteine insgesamt nur ca. 100 Lautelemente benötigt werden. Bei der Aufbereitung der Lautelemente, also in der Analysephase, sind die natürlichen Sprachproben, aus denen die zu verwendenden Lautelemente gewonnen werden, ohnehin zu untersuchen, beispielsweise die oben erwähnten "unempfindlichen" Stellen zu bestimmen. Dabei lassen sich diese Sprachproben rechnerisch modifizieren, insbesondere bei Übergangslauten Diskontinuitäten in den Formantverläufen eleminieren.

Die Ausnutzung von Lauttransformationen, d.h. einer gegenseitigen Überführbarkeit verwandter Laute, war bereits Gegenstand beim aus der DE-OS 25 31 006 bekannten Stand der Technik, von dem die Erfindung ausgeht. Dort führte die Reduzierung z.B. bei den Konsonanten von 22 auf 8. Weiterhin waren etliche Ausnahmen, etwa 150 Übergänge, je eine Pitchperiode stimmhafter Laute sowie ein Abschnitt aus dem Mittelteil der stimmlosen Laute und schließlich bei Explosivlauten noch der Anfang der Zeitfunktion zu speichern. Bei der Erfindung ergibt sich eine erhebliche Reduzierung aufgrund folgender Maßnahmen: Übergänge - ausgenommen Plosivlautkombinationen - lassen sich zeitlich invertieren; durch Verlängern bzw. Verkürzen der Lautdauer finden Vokalumwandlungen statt, durch Verkürzen der Lautdauer ergeben sich auch Konsonantenumwandlungen. Die benötigten Lautelemente setzen sich dadurch zusammen aus knapp 60 Elementen für Übergangslaute, 27 Elementen für stimmhafte Einzellaute und 13 Elementen für stimmlose Einzellaute. Weitere Einzelheiten dazu folgen noch im Zusammenhang mit der Figurenbeschreibung.

Besonders bevorzugte Ausführungsformen der Erfindung bestehen darin, in den digital gespeicherten Elementen für die stimmhaften Einzellaute zum Zwecke der Tonhöhenvariation zusätzliche Abtastwerte vorzusehen. Diese Maßnahme führt

zwar zu einer geringfügigen Erhöhung um ca. 1000 Byte des benötigten Speicherplatzvolumens, ermöglicht aber weitergehende Variationen in den Melodieverläufen.

Im engen Zusammenhang damit ist es weiterhin vorteilhaft, wenn ein zusätzlicher Abtastwert einen zwischen den benachbarten wahren Abtastwerten liegenden interpolierten Wert besitzt. Auf diese Weise lassen sich eventuelle Diskontinuitäten verringern oder vermeiden, die zwischen den wahren Abtastwerten, die auf jeden Fall benötigt und verwendet werden, auftreten würden.

Wie bereits weiter oben schon erwähnt, sind für die Maßnahmen zur Melodievariation "unempfindliche" Stellen in den Zeitverläufen bevorzugt, d.h. Markierwörter sind vorzugsweise an Stellen geringer Steigung des Zeitsignals vorzusehen. Ein zugehöriges Fehlersignal weist an solchen Stellen sehr kleine Ausschläge auf und erlaubt damit auf einfache Weise, die gewünschten diskreten Stellen zu ermitteln, zu lokalisieren und zu markieren.

Manchmal, besonders bei großen, erwünschten Tonhöhenschwankungen, kann es erforderlich sein, den möglichen Bereich der für Auslassungen bzw. die Verwendung geeigneten Abtastwerte voll auszunutzen. Häufiger sind jedoch die Fälle, in denen nur einige der zur Verfügung stehenden vorgegebenen Abtastwerte benötigt werden. Aus diesem Grunde ist es günstig, wenn Markierwörter an Stellen geringer Steigung des Zeitsignals mit einer höheren Priorität für Tonhöhenvariation ausgestattet sind als solche an Stellen mit größerer Steigung. Das bedeutet, zunächst erfolgen derartige Veränderungen immer an den unempfindlichsten Stellen, gegebenenfalls werden aber auch die empfindlicheren Stellen dazu herangezogen.

Obwohl durchaus auch die Möglichkeit besteht, bei den für Tonhöhenvariation als geeignet vorgegebenen Abtastwerten

getrennt vom gespeicherten Lautelement die zugehörigen Adressen zu verwalten, wird bei den Ausführungsformen der Erfindung die Lösung mit den Markierwörtern bevorzugt. Dabei können ein Markierwort und ein wahrer oder zusätzlicher Abtastwert digitale Muster desselben Vorrats aufweisen. Hinsichtlich einer eindeutigen Unterscheidbarkeit zwischen Markierwort und Abtastwert sollen dann jedoch Markierwörtern digitale Muster vorbehalten sein, die bei den Abtastwerten nicht vorkommen.

Allein schon aus Gründen unterschiedlicher Prioritäten reicht ein einziges Muster für Markierwörter nicht aus. Da eine softwaremässige Identifizierung der Muster keine besondere Systematik bei der Verteilung der digitalen Muster erfordert, ist es ohne weiteres möglich, für Markierwörter die Muster mit den höchsten Stellenzahlen, bei 8-bit-Wörtern z.B. die Muster 246, 247, ... 255, vorzubehalten. Diese Muster können bei der Digitalisierung der Abtastwerte deshalb auf besonders vorteilhafte Weise ausgespart werden, weil eine Begrenzung am oberen Ende zu kaum spürbaren Beschränkungen führt.

Von besonderer Bedeutung ist es für Ausführungsformen der Erfindung, während der Wortpausen die Gestalt der für die Verkettung des nächstfolgenden Wortes benötigten Lautelemente anhand der Eingangsbefehle bestimmen zu können. Hierdurch werden Diskontinuitäten bei der Ausgabe der einzelnen Wörter vermieden. Die Dauer für die Bestimmung der Gestalt der benötigten Synthesebausteine liegt, auch für sehr lange Wörter, im Bereich von wenigen Millisekunden. Unter Bestimmung der Gestalt ist hier zu verstehen: aufsuchen des betreffenden Lautelements, gegebenenfalls zeitlich invertieren, Lautdauer verlängern bzw. verkürzen und Wiederholanzahl des gespeicherten Lautelements angeben.

Ein weiterer wesentlicher Vorzug der Erfindung besteht darin, daß über eine alphanumerische Tastatur eingegebene Folgen üblicher Schriftzeichen in einem dem eigentlichen Synthesevorgang vorausgehenden Verfahrensschritt selbsttätig in eine als Eigangsbefehle geeignete Folge von Lautschriftzeichen transkribiert werden kann. Hierdurch wird auch ungeübten bzw. nicht geschulten Benutzern die Anwendung erheblich erleichtert bzw. überhaupt erst eröffnet. Selbstverständlich bleibt dabei auch die Möglichkeit bestehen, Lautschriftzeichen bzw. die geeigneten Eingangsbefehle unmittelbar einzugeben.

Für die Transkription ist allerdings weiteres Speichervolumen erforderlich. Überraschend ist dabei, daß dafür jedoch nur etwa ein Drittel desjenigen Speicherplatzvolumens benötigt wird, der für die Synthese vorzusehen ist, d.h. etwa ein Viertel des gesamten Speicherplatzvolumens für Synthese und Transkription, wenn die Transkription auf folgende Art durchgeführt wird: zunächst werden lexikalisch erfaßte Ausnahmen und Fremdwörter bearbeitet; ansonsten wird der Wortschatz einer Präfixverarbeitung, unter Berücksichtigung von Ausnahmen, einer Endungsabspaltung und einer Suffixverarbeitung, ebenfalls unter Berücksichtigung von Ausnahmen, unterzogen und die Transkription der Wortstämme nach katalogartig gespeicherten Regeln durchgeführt. Diese oder ähnliche Maßnahmen sind für Sprachwissenschaftler an sich geläufig.

Eine Schaltungsanordnung zur Durchführung des erfindungsgemäßen Verfahrens kann mit einem Mikroprozessor aufgebaut sein, an den Festwertspeicher mit einer Speicherkapazität von insgesamt 32 kByte und ein Arbeitsspeicher für 1 kByte anzuschließen sind, und weist außerdem einen dekompandierenden Digital-Analog-Wandler und einen -lautstärkeregelbaren- Niederfrequenzverstärker und einen Lautsprecher als elektro-akustische Wandlereinrichtung auf. Derartige Schaltungselemente und Bauteile sind marktüblich. Das Konzept

ermöglicht aber auch eine weitgehende Integration. Die Dekomparadierung vor der Digital-Analog-Wandlung beinhaltet selbstverständlich, daß zuvor die gespeicherten Daten einer die Datenrate reduzierenden Kodierung unterzogen wurden. Gebräuchliche und in der angegebenen Reihenfolge immer stärker reduzierende Verfahren sind die logarithmische PCM und die Adaptive-Delta-PCM. Aus gebräuchlichen Sprach-Übertragungssystemen sind betreffende Bauteile bekannt und ohne weiteres auch bei Ausführungsformen der Erfindung einzusetzen.

Hinsichtlich des Aufwandes bei Schaltungsanordnungen sind noch immer die Speicher, genauer gesagt deren Größe, von Bedeutung. Deshalb ist es wichtig für Kostenabschätzungen, daß bei einer Schaltungsanordnung zur Durchführung des erfindungsgemäßen Verfahrens die Aufteilung der Kapazität von Festwertspeichern in:

1,5 kByte für das Transkriptionsprogramm,
6 kByte für die Transkriptionsgrammatik,
1,5 kByte für das Syntheseprogramm,
1 kByte für die Synthesematrix
und 22 kByte für die Lautelemente
erfolgen kann.

Schließlich ist es für die verschiedenartigen Einsatzgebiete von Ausführungsformen der Erfindung wichtig, daß die Eingabe der Daten, d.h. der Schreib- oder Lautschriftsymbolfolgen, sowie die Ausgabe der akustischen Signale sowohl direkt am Gerät als auch jeweils an entfernten Orten erfolgen kann. Dazu kann entsprechend am Eingang z.B. eine V24-Schnittstelle bzw. am Ausgang eine Niederfrequenzbuchse vorgesehen sein.

Die Anwendungsmöglichkeiten für ein derartiges Sprachsynthesesystem sind aufgrund der Möglichkeit, ein unbegrenztes Vokalubar zu generieren, äußerst mannigfaltig. Beispielhaft sollen erwähnt sein: Telefon-Auskunftssysteme; akustischer

Ersatz oder Unterstützung bei unübersichtlichen Anzeigetafeln, insbesondere Flug- oder Fahrplänen; Ersatz oder Ergänzung dort, wo die Aufmerksamkeit von Personen durch Dauerbeobachtung einzelner Ziffern- oder Textanzeigen oder Warnanlagen über Gebühr beansprucht wird, z.B. bei Flugzeug-Bordsystemen; Tastenwahltelefone als Eingabetastatur und Telefonhörer als Ausgabe bei Datenverarbeitungsanlagen, z.B. für Auskünfte sich laufend ändernder Daten, wie Lagerbestände, Kontenstände, Börsenkurse, medizinische Diagnosen oder laufende Überwachung von Körperfunktionen von Patienten im Krankenhaus oder zu Hause; Bestellungen von Waren nach Katalognummern, von Theater- oder Konzertkarten; Erteilung und Annahme von Aufträgen, Umdispositionen u. dgl.; Fernübertragung von Prozessdaten; Hausleitsysteme; Sprachen-Unterricht; Computergestützter Unterricht; Verkehrsleitung; Bibliotheken-Anfragen und Auskünfte; Lexikon- Auskunftsdienst, Hilfe für Behinderte -Sprach- und Sehbehinderte- und vieles mehr.

In den Zeichnungen sind Einzelheiten für Ausführungsformen der Erfindung schematisch dargestellt. Dabei zeigen:

Fig. 1: ein Blockschaltbild für ein Sprachsynthesegerät mit Transkriptionseinheit,

Fig. 2: ein Blockschaltbild eines Sprachsynthesegerätes mit Transkriptionseinheit, auf Mikroprozessorbasis;

Fig. 3: eine Darstellung der Lage der drei ersten Formanten für verschiedene Laute;

Fig. 4: eine Darstellung von Formantsprüngen an den Übergängen zwischen drei Einzellauten;

Fig. 5: eine Darstellung für die Reduktionsmöglichkeit der Länge von Elementen;

Fig.6a: ein Beispiel für zeitliche Invertierung von Übergangslauten;

Fig.6b: die Möglichkeiten für Vokalumwandlungen;

Fig.6c: die Möglichkeiten für Konsonantenumwandlungen;

Fig. 7: ein Beispiel für die Veränderung des Hör-eindrucks durch Verschieben des Anfangs-punktes;

Fig. ein Beispiel für die rechnerische Modifi-zierung

8a,b,c: eines stimmhaften Einzellautes zur Varia-tion der Tonhöhe;

Fig. 9: ein -auszugsweises- Beispiel für die An-ordnung von wahren, auslaßbaren und zu-sätzlichen Abtastwerten sowie von Mar-kierwörtern in einem gespeicherten Ele-ment eines stimmhaften Einzellautes;

Fig. 10: eine Darstellung der Aufteilung und des Inhaltes des Lautelemente-Speichers;

Fig. 11: eine Darstellung des Ablaufs einer Trans-kription

und Fig. 12: eine Darstellung eines Synthesebeispiels (monoton).

Wie die Fig. 1 zeigt, besteht eine Sprachsynthesesystem bei Ausführungsformen nach der Erfindung im wesentlichen aus zwei Einheiten, der für die Transkription und der für die Synthese selbst. Einzugeben ist entweder eine Schriftzei-chenfolge, was über eine alphanumerische Tastatur oder über eine V24-Schnittstelle geschehen kann, oder aber eine Laut-zeichenfolge. Obwohl geübte bzw. geschulte Benutzer über geeignete Tastaturen auch die Lautzeichenfolgen unmittelbar eingeben können, wird in den meisten Anwendungsfällen bei einem Verzicht auf die Transkription die Syntheseeinheit dann wohl die entsprechenden Eingangssignale von einem ent-fernten Ort über eine Datenleitung und die V24- Schnitt-stelle erhalten. Selbstverständlich lassen sich auch andere Schnittstellenbedingungen einhalten und im Rahmen fachmän-nischen Könnens realisieren. Die Transkriptionseinheit greift auf vorbereitete Regeln, unter dem Begriff Grammatik zusammengefaßt, zurück, die Syntheseeinheit im wesentlichen auf die gespeicherten Lautelemente. Die synthetisierten

Abtastwertfolgen gelangen über einen Digital-Analog-Wandler D/A und einen regelbaren Verstärker entweder direkt über einen Lautsprecher oder über eine Niederfrequenzbuchse und eine nicht dargestellt Sprachübertragungsleitung und am entfernten Ort über einen Lautsprecher als Schallwellen zur Wieder-, besser Ausgabe.

Das in Fig. 2 dargestellte Blockschaltbild gibt insbesondere im Größenvergleich der einzelnen Blöcke den Speicherplatzbedarf mit den Anteilen wieder, die für die Synthese und die Transkription insgesamt benötigt werden. Das System ist auf Basis eines Mikroprozessors µP konzipiert. Für die Eingabe der Schriftzeichenfolgen ist eine alphanumerische Tastatur, für die Ausgabe der als Sprache wahrnehmbaren Schallwellen ein üblicher elektro-akustischer Wandler vorgesehen. Für die Transkription arbeitet der Mikroprozessor µP mit dem Transkriptionsprogramm TP und der Transkriptionsgrammatik TG, bei der Sprachsynthese mit dem Syntheseprogramm SP und der Synthesematrix SM, wobei die benötigten Lautelemente je nach Bedarf aus dem Lautelementespeicher SE entnommen, in die im Arbeitsspeicher RAM abgelegte, aus der betreffenden Lautzeichenfolge abgeleitete Gestalt gebracht, in der betreffenden Anzahl und Reihenfolge verkettet und an den Digital-Analog-Wandler (s. Fig. 1, D/A) übergeben werden. Eine Lautstärkeregelung innerhalb der synthetisierten Wörter und Sätze erfolgt, ebenfalls vom Mikroprozessor µP gesteuert und entsprechend dafür eingegebener Befehle, im regelbaren Niederfrequenzverstärker (s. Fig. 1) vor der Abstrahlung der Schallwellen bzw. der Übertragung des Niederfrequenzsignals.

Die in Fig. 3 dargestellte Lage der drei ersten Formanten für neun verschiedene Laute läßt erkennen, daß insbesondere der erste und der zweite Formant von erheblicher Bedeutung für die Lautbildung sind. Aufgrund der linearen Teilung der Frequenzskala darf jedoch nicht übersehen werden, daß auch

beim dritten Formanten der Bereich etwa einer halben Oktave beansprucht wird.

In Fig. 4 ist für drei Laute die Lage der Formanten dargestellt. Es zeigt sich, daß an den Übergängen teilweise recht erhebliche Sprünge auftreten, die als äußerst unangenehm wahrgenommen werden würden. Hierbei handelt es sich jedoch um bekannte Erscheinungen, die lediglich deshalb nicht unerwähnt bleiben sollen, um die Vielschichtigkeit der Probleme anzudeuten, die bei einem Sprachsynthesesystem zu beachten sind.

Das in Fig. 5 dargestellte Zeitsignal des Wortes "Asche" soll die Möglichkeit der Reduktion der Länge von Lautelementen durch Segmentierung in quasistationäre Bereiche S und Übergangsbereiche Ü veranschaulichen. Innerhalb der quasistationären Bereiche S sind Sprachgrundfrequenzperioden P zu erkennen, die den signifikanten Bereich eines Lautes bilden und nur in dieser Länge als Element für die Synthese abgespeichert zu werden brauchen. Ähnliche Grundfrequenzperioden sind auch bei Übergangsbereichen zu erkennen und reichen als Synthesebaustein ebenfalls aus.

Die in den Fig. 6a, 6b und 6c angegebenen Möglichkeiten für zeitliche Invertierung von Übergängen (Fig. 6a), für Vokalumwandlung (Fig. 6b) und für Konsonantenumwandlung (Fig. 6c) sprechen für sich und bedürfen deshalb hier keiner näheren Erläuterung. Allerdings ist, wie weiter oben bereits erwähnt, darauf hinzuweisen, daß eine Verkürzung oder Verlängerung der Lautdauer eben nicht nur eine Verlagerung der Tonhöhe mit sich bringt, sondern insbesondere eine Lautumwandlung bewirkt. Von den 16 in Fig. 6 c angegebenen Lauten brauchen übrigens nur die in jeder Zeile an erster Stelle angegebenen gespeichert zu werden. Dies sind zwar die Laute mit den jeweils meisten benötigten Abtastwerten, doch wird dadurch Speicherplatz von gut 60 % gegenüber einer Speicherung aller dieser Laute eingespart.

Die in Fig. 7 dargestellte Veränderung des Höreindrucks gibt an, daß 20 Testpersonen eine Konsonantenumwandlung feststellen sollten (in Klammern), die - bis auf zwei Personen bei der Verschiebung des Anfangspunktes auf 160 ms - den angegebenen Höreindruck bei den einzelnen Umwandlungsformen bestätigten.

Die Fig. 8a, 8b und 8c zeigen an einem Beispiel, auf welche Weise die bei der Erfindung wesentliche Variation der Tonhöhe ermöglicht wird. In Fig. 8a ist eine Grundfrequenzperiode des Lautes /a/ aufgetragen. Zur Modifizierung wird zunächst von einem Prädiktionsfehlerfilter das dazugehörige Fehlersignal (Fig. 8b) erzeugt. Daraus ist zu erkennen, daß diskrete Stellen angegeben werden können, an denen Modifizierungen vorzunehmen sind, ohne den Lautcharakter, jedoch seine Tonhöhe zu verändern. In Fig. 8c ist die gegenüber Fig. 8a um etwa 20 % gekürzte Periode des Lautes /a/ angegeben. Es zeigt sich im Vergleich der Kurvenverläufe von Fig. 8a und 8c, daß eine Verkürzung der Periode, d.h. eine Erhöhung der Tonhöhe, das eigentliche charakteristische Bild nicht verändert, der Laut /a/ als solcher also erhalten bleibt und -wie gewünscht- höher klingt.

In der Fig. 9 ist ein Beispiel - auszugsweise - angegeben, in welcher Reihenfolge (lfd. Nr.) in einem gespeicherten Element eines in der Tonhöhe veränderbaren, stimmhaften Übergangs- oder Einzellautes wahre Abtastwerte WAW, auslaßbare Abtastwerte DAW, zusätzliche Abtastwerte ZAW und Markierwörter MAW aufeinanderfolgen. Im Normalfall, d.h. wenn keine Tonhöhenvariation erfolgen soll, werden nur die wahren Abtastwerte WAW verwendet. Für eine Absenkung der Tonhöhe werden zusätzliche Abtastwerte ZAW mit verwendet, für eine Erhöhung hingegen gegenüber dem Normalfall auslaßbare Abtastwerte DAW weggelassen. Mit den Markierwörtern werden nicht nur die zusätzlichen ZAW bzw. auslaßbaren Abtastwerte DAW lokalisiert, sondern vorteilhaft auch deren Priorität für Tonhöhenänderungen bestimmt.

Der in Fig. 10 dargestellte Block soll das Verhältnis des Speicherplatzbedarfs veranschaulichen, der für die Synthesebausteine, die Elemente der Einzel- und der Übergangslaute, benötigt wird. Dabei handelt es sich in erster Linie um die wahren Abtastwerte WAW der Elemente, außerdem aber auch um die Markierwörter MAW und die rechnerisch bestimmten zusätzlichen Abtastwerte ZAW bei den stimmhaften Einzellauten bzw. den stimmhaften Bereichen von Übergangslauten. Die gestrichelte Linie zwischen den Bereichen für die Einzellaut- und die Übergangslaut-Elemente zeigt eine Aufteilung etwa im Verhältnis 4 : 6.

Die Fig. 11, in der der Ablauf einer Transkription dargestellt ist, spricht für sich, soll aber anhand eines Beispiels, die Transkription des Wortes "verwischend" näher erläutert werden:

Bei der lexikalischen Verarbeitung ergibt sich, daß es sich um keine Ausnahme handelt. Die Wortanalyse erfolgt also nach:

Präfix: "ver"
Stamm: "wisch"
Suffix: "en"
Endung: "d"

Bei der Transkription des Stammes nach Regeln ist festzustellen, ob die Aussprache der Symbolfolge "sch" als ein Laut /sch/ (wie in: Schule) oder als zwei getrennte Laute /s/ und /ch/ erfolgen muß. Dazu gelten folgende Regeln aus dem Katalog: Befinden sich vor "sch" zwei Vokale oder ein Umlaut, gilt zunächst die zweite Alternative, also zwei getrennte Laute /s/ und /ch/ (Beispiel: Röschen/Roeschen). Ist dabei jedoch der zweite Vokal ein "u", gilt dennoch die erste Alternative, d.h. der Einzellaut /sch/ (Beispiel: tauschen).
Befinden sich vor "schen" drei Vokale, wobei ein Umlaut wiederum als zwei Vokale angesehen werden, gilt wieder die

zweite Alternative, also zwei getrennte Laute /s/ und /ch/ (Beispiel: Häuschen/Haeuschen). Ausnahmen hiervon sind nur zwei Wörter: täuschen/taeuschen und Geräuschen/Geraeuschen.

Ein weiteres Beispiel aus dem umfangreichen Regelkatalog betrifft den Laut /ch/. Dabei werden unterschieden:
"euch" /s-euch-e, f-euch-t/,
"uch" /a-uch, s-uch-t, gebra-uch-en, anspr-uch/,
"och" /n-och, t-och-ter, h-och, denn-och/,
"ach" /n-ach, d-ach, gem-ach-t, spr-ach-, -ach-t/,
"ch" /-ch-arakter/,
"ch" /ni-ch-t, dur-ch, wel-ch-es, re-ch-t/,
wobei hier jeweils nur einige Lautbeispiele aufgeführt sind.

Die Fig. 12 zeigt den Signalverlauf - monoton - des synthetisierten Wortes /Tasche/. (Eine den Signalverlauf, die Melodik, Rhythmik und Dynamik enthaltende Darstellung wäre, soweit mit gebräuchlichen Mitteln überhaupt möglich, zweifellos unübersichtlicher). Für das /t/ wurde ein gekürztes /s/ verwendet. Der Übergang /ta/ entstammt dem Doppellaut /sa/. Für das /a/ wurden einer Periode 8 Wiederholungen angefügt. Der Übergang /asch/ wurde dem Doppellaut /sa/, zeitlich invertiert, entnommen. Beim /sch/ handelt es sich um einen stimmlosen Einzellaut. Der Übergang /schə/ entstammt dem Doppellaut /sə /. Schließlich wurde für das /ə / am Ende zunächst eine Periode 6mal und sodann noch 6mal, jedoch mit dem Ausschnitt einer Sinusfunktion bewertet, wiederholt.

HEINRICH-HERTZ-INSTITUT FÜR NACHRICHTENTECHNIK BERLIN GMBH

Patentansprüche

1. Verfahren zur Synthese von Sprache mit unbegrenztem Wortschatz im Zeitbereich aus Lautelementen, die aus natürlichen Sprachproben gewonnen und in digitaler Form, redundanzarm kodiert, gespeichert und außerdem im Hinblick auf den erforderlichen Speicherplatzbedarf in der Länge jeweils auf den signifikanten Bereich des betreffenden lauttypischen Zeitsignals und in der Anzahl unter Ausnutzung sich gegenseitig ineinander überführbarer verwandter Laute reduziert sind, wobei zur Sprachsynthese diese Lautelemente aufgrund von Eingangsbefehlen und von vorgegebenen Verknüpfungsregeln in der erforderlichen Gestalt, Anzahl und Reihenfolge zu digitalen Signalfolgen verkettet werden, aus denen mittels Digital-Analog-Wandlung und steuerbarer Verstärkung als Sprache wahrnehmbare Schallwellen erzeugt werden, d a d u r c h g e k e n n z e i c h n e t , daß insgesamt ca. 100 Lautelemente, nämlich:

   - etwa 50 Elemente für Übergangslaute mit je durchschnittlich 240 Abtastwerten für 8 kHz Ausgabefrequenz

   und

   - etwa 40 Elemente für Einzellaute mit je durchschnittlich 500 Abtastwerten bei stimmlosen und 140 Abtastwerten bei stimmhaften Einzellauten und 8 kHz Ausgabefrequenz vorgesehen sind,

   und daß die Tonhöhe für die Wiedergabe bei den Elementen für die stimmhaften Übergangs- und Einzellaute veränderbar ist, indem solche Abtastwerte, die an diskreten Stellen des Zeitsignals mittels Markierwörtern als geeignet vorgegeben sind, je nach Bedarf aufgrund entsprechender Eingangsbefehle bei der Bildung der digitalen Signalfolgen ausgelassen bzw. mindestens einmal verwendet werden.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß in den digital gespeicherten Elementen für die stimmhaften Laute zum Zwecke der Tonhöhenvariation zusätzliche Abtastwerte vorgesehen sind.

3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, daß ein zusätzlicher Abtastwert einen zwischen den benachbarten wahren Abtastwerten liegenden interpolierten Wert besitzt.

4. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß Markierwörter vorzugsweise an Stellen geringer Steigung des Zeitsignals vorgesehen sind.

5. Verfahren nach Anspruch 4, dadurch gekennzeichnet, daß Markierwörter an Stellen geringerer Steigung des Zeitsignals mit einer höheren Priorität für Tonhöhenvariation ausgestattet sind als solche an Stellen mit größerer Steigung.

6. Verfahren nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß Markierwörtern digitale Muster vorbehalten sind, die bei den Abtastwerten nicht vorkommen.

7. Verfahren nach Anspruch 6, dadurch gekennzeichnet, daß für Markierwörter die Muster mit den höchsten Stellenzahlen, bei 8-bit-Worten z.B. die Muster 246, 247, ... 255, vorbehalten sind.

8. Verfahren nach einem der Ansprüche 1 bis 7, dadurch gekennzeichnet, daß während der Wortpausen die Gestalt der für die Verkettung des nächstfolgenden Wortes benötigten Lautelemente anhand der Eingangsbefehle bestimmt wird.

9. Verfahren nach einem der Ansprüche 1 bis 8, dadurch gekennzeichnet, daß über eine alphanumerische Tastastur eingegebene Folgen üblicher Schriftzeichen in einem dem eigentlichen Sprachsynthesevorgang vorausgehenden Verfahrensschritt selbsttätig in eine als Eingangsbefehle geeignete Folge von Lautschriftzeichen transkribiert wird.

10. Verfahren nach Anspruch 9, dadurch gekennzeichnet, daß zunächst lexikalisch erfaßte Ausnahmen und Fremdwörter bearbeitet werden, und der Wortschatz ansonsten einer Präfixverarbeitung, unter Berücksichtigung von Ausnahmen, einer Endungsabspaltung und einer Suffixverarbeitung, ebenfalls unter Berücksichtigung von Ausnahmen, unterzogen wird, und die Transkription der Wortstämme nach katalogartig gespeicherten Regeln durchgeführt wird.

11. Schaltungsanordnung zur Durchführung des Verfahrens nach einem der Ansprüche 1 bis 10, gekennzeichnet durch einen Mikroprozessor (µP), an den Festwertspeicher (ROM) mit einer Speicherkapazität von insgesamt 32 kByte und ein Arbeitsspeicher (RAM) für 1 kByte angeschlossen sind, sowie durch eine an sich bekannte, aus einem dekompandierenden Digital-Analog- Wandler und einem Niederfrequenzverstärker und einem Lautsprecher bestehende elektro-akustische Wandlereinrichtung.

12. Schaltungsanordnung nach Anspruch 11, gekennzeichnet durch eine Aufteilung der Kapazität der Festwertspeicher (ROM) in: 1,5 kByte für das Transkriptionsprogramm,
6 kByte für die Transkriptionsgrammatik,
1,5 kByte für das Syntheseprogramm,
1 kByte für die Synthesematrix und
22 kByte für die Lautelemente.

13. Schaltungsanordnung nach Anspruch 11 oder 12, gekennzeichnet durch eine V24-Schnittstelle am Eingang.

14. Schaltungsanordnung nach einem der Ansprüche 11 bis 13, gekennzeichnet durch eine Niederfrequenzbuchse am Ausgang.

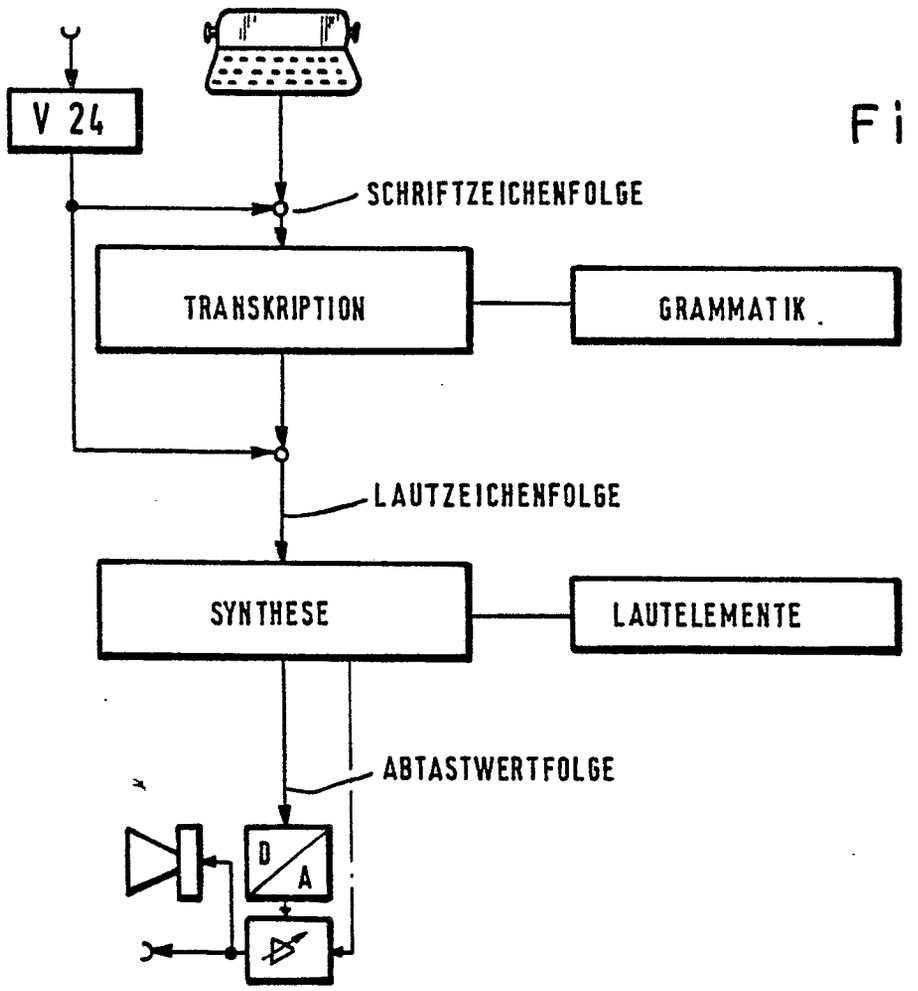

V 24
SCHRIFTZEICHENFOLGE
TRANSKRIPTION
GRAMMATIK
LAUTZEICHENFOLGE
SYNTHESE
LAUTELEMENTE
ABTASTWERTFOLGE
D
A

Fig.1

Fig.2

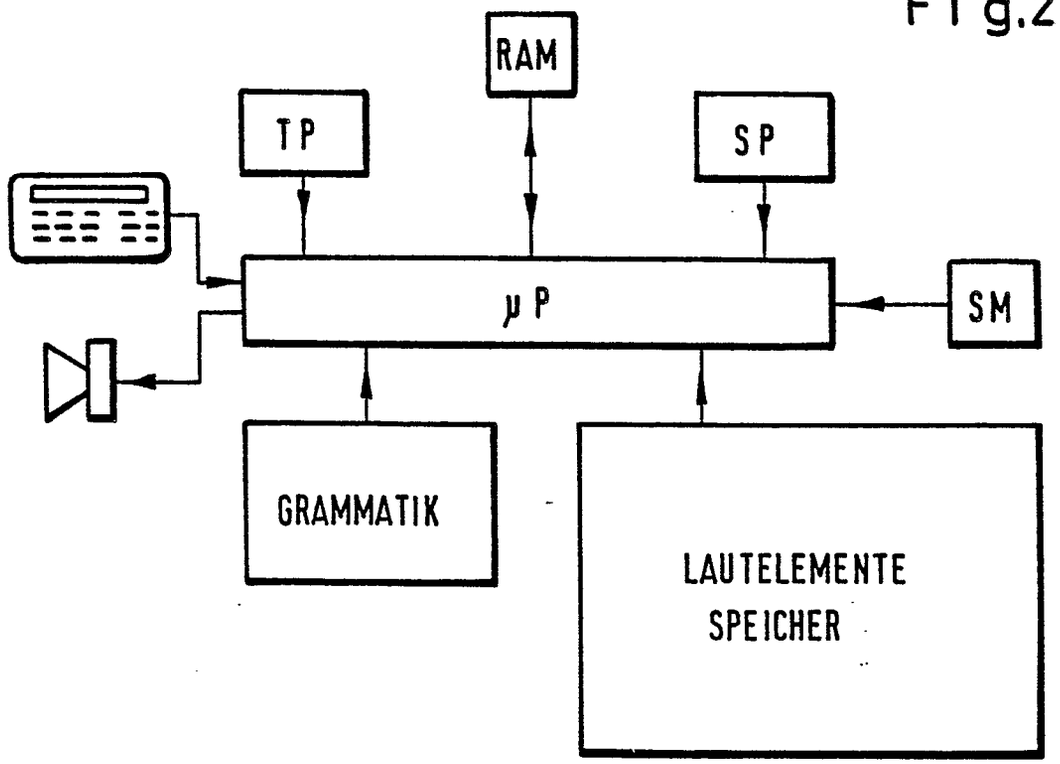

RAM
TP
SP
μP
SM
GRAMMATIK
LAUTELEMENTE
SPEICHER

Fig.3

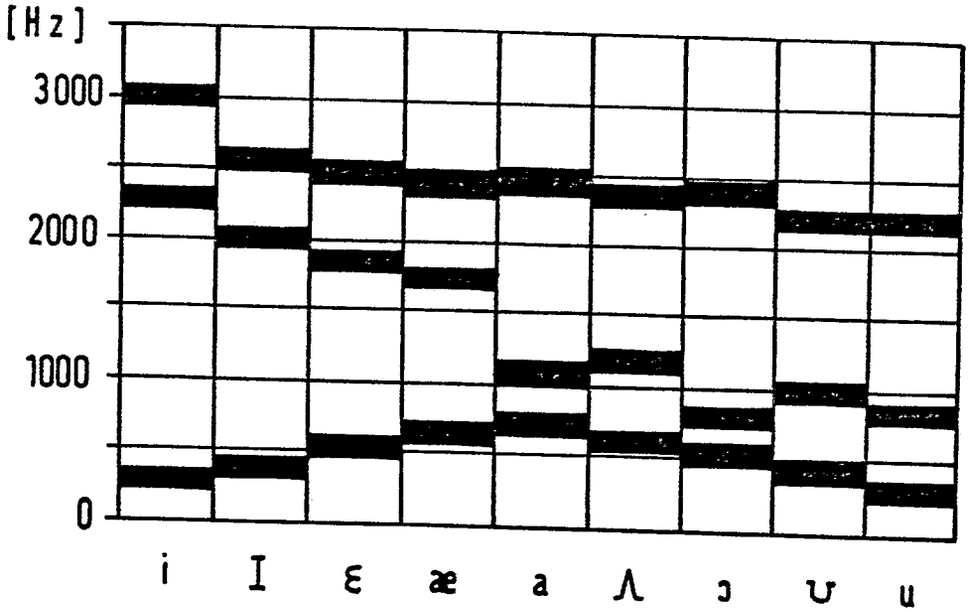
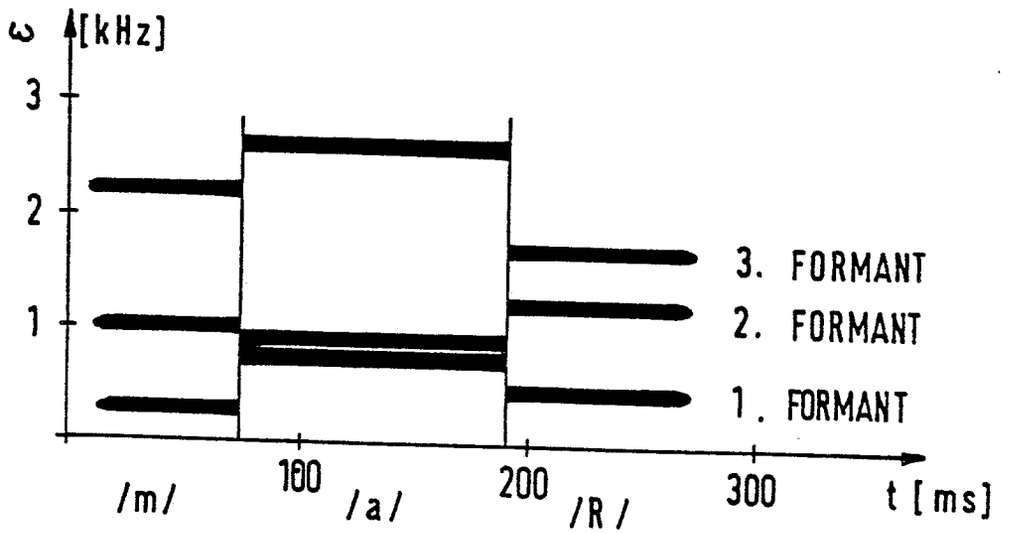

[Hz]
3000
2000
1000
0
i
I
ɛ
æ
a
ʌ
ɔ
ʊ
u

Fig.4


ω
[kHz]
3
2
1
3. FORMANT
2. FORMANT
1. FORMANT
/m/
100
/a/
200
/R/
300
t [ms]

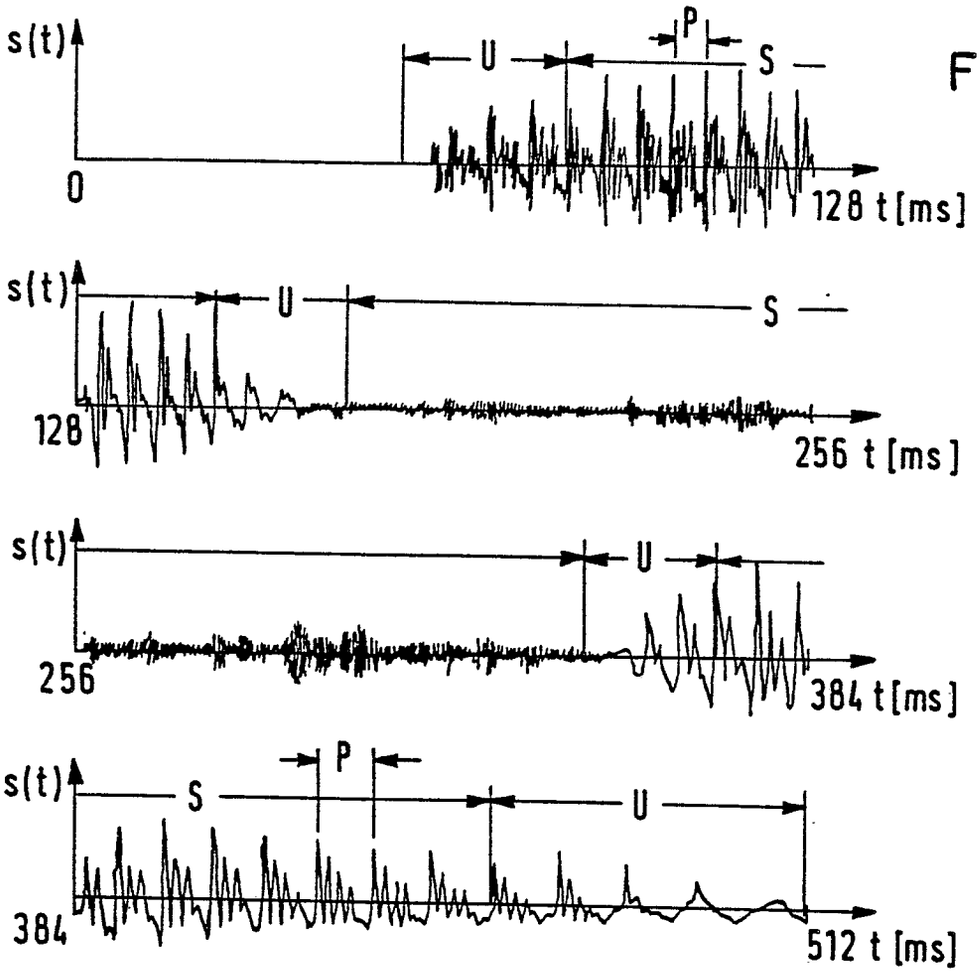


Fig.5

Fig.6a

/mo/ ⟷ /om/

/a/ ⟷ /a:/
/I/ ⟷ /e/
/ʊ/ ⟷ /o/
/ɛ/ ⟷ /ɛ:/
/y/ ⟷ /ø/

Fig.6b

/ʃ/ → /tʃ/ → /t/ → /d/
/s/ → /ts/ → /t/ → /d/
/z/ → /t/ → /d/
/v/ → /b/
/f/ → /pf/ → /p/

Fig.6c

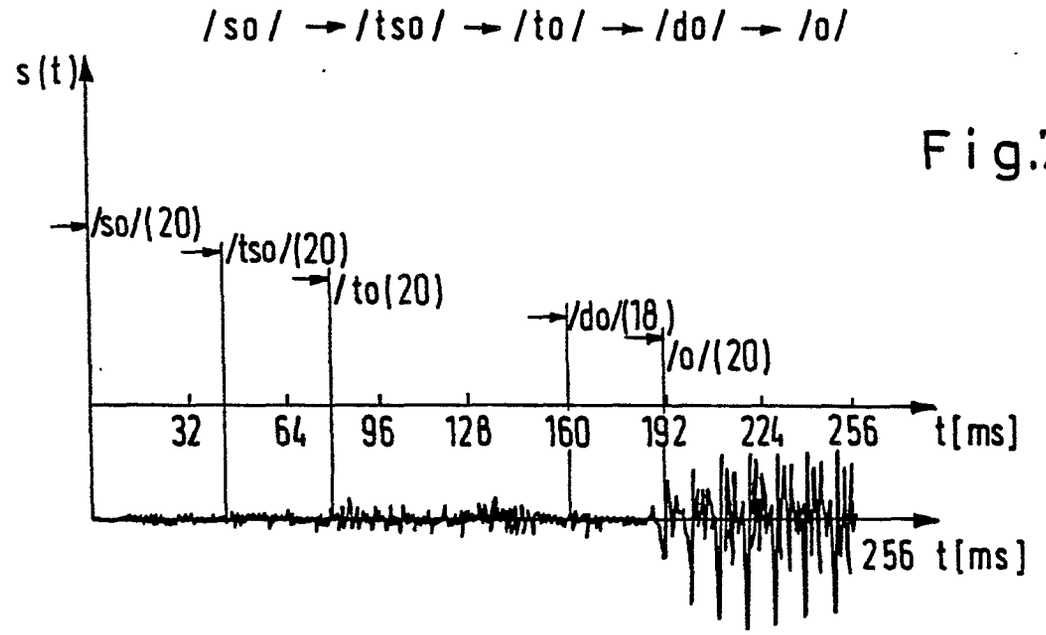


Fig.7

Fig.8a

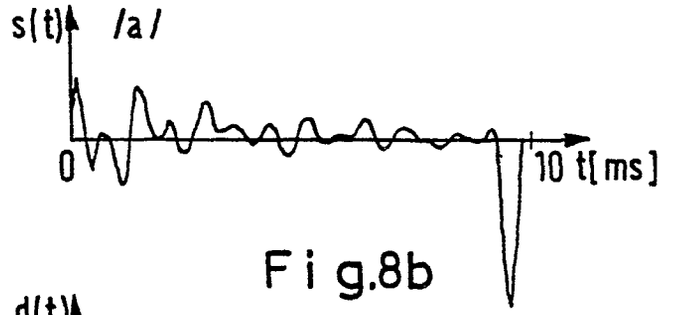
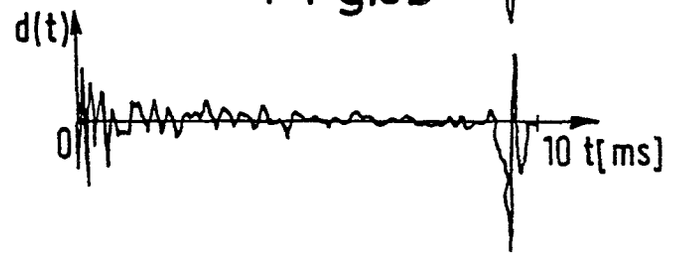


Fig.8b



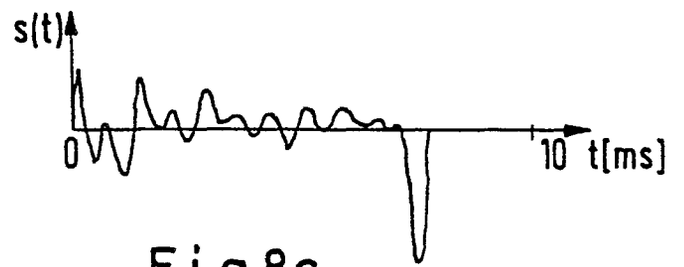


Fig.8c

Lfd. Nr.	WAW	DAW	MAW	ZAW
1	×			
2	×			
3	×			
4	×			
5	×			
⋮				
29	×	×		
30			×	
31				×
⋮				
55	×	×		
56			×	
57				×
58	×			
⋮				
135	×			
136	×	×		
137			×	
138				×
139	×			
140	×			

Fig.9

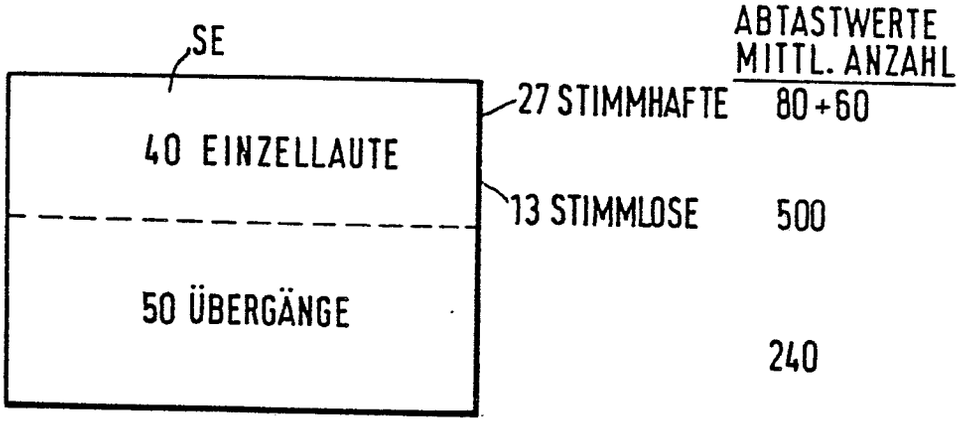


Fig.10

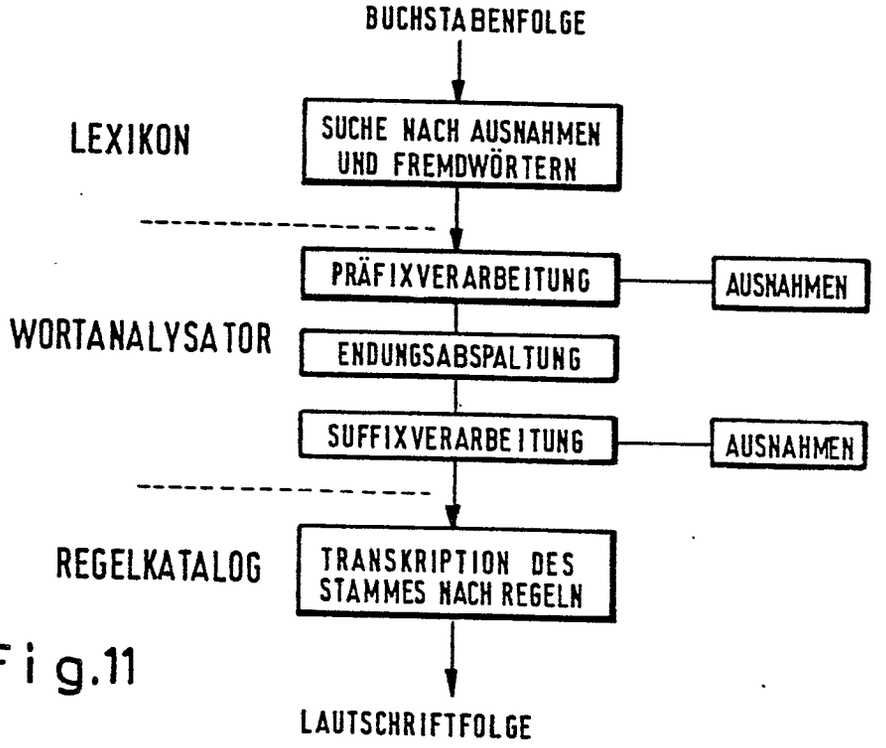


Fig.11

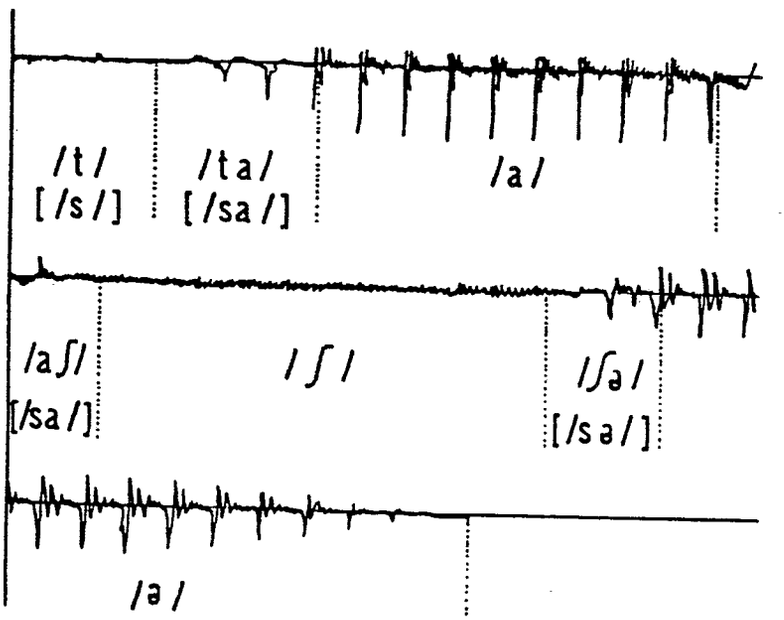


Fig.12