

# 公告本

申請日期	88. 4. 28
案 號	88106840
類 別	G10L 5/04

A4  
C4

422967

(以上各欄由本局填註)

## 發明專利說明書

一、發明名稱	中 文	利用決定樹來產生及評定拼字的多重發音的方法與裝置
	英 文	METHOD AND APPARATUS USING DECISION TREES TO GENERATE AND SCORE MULTIPLE PRONUNCIATIONS FOR A SPELLED WORD
二、發明人	姓 名	(1)羅朗·庫 (2)珍克勞德·瓊魁 (3)馬提歐·坎東里尼
	國 籍	(1)加拿大(2)法國 (3)義大利
	住、居所	(1)美國加州聖芭巴拉·凱樂席塔3928號 (2)美國加州聖芭巴拉·聖大阿那路146號 (3)美國加州聖芭巴拉·克利夫道821號#B-1
三、申請人	姓 名 (名稱)	日商·松下電器產業股份有限公司
	國 籍	日本
	住、居所 (事務所)	日本國大阪府門真市大字門真1006番地
	代 表 人 姓 名	森下洋一

裝

訂

線

(由本局填寫)

承辦人代碼：
大 類：
IPC分類：

A6  
B6

本案已向：

美 國 ( 地 區 )	申 請 專 利 , 申 請 日 期 :	案 號 :	<input checked="" type="checkbox"/> 有 <input type="checkbox"/> 無 主 張 優 先 權
	1998,4,29	09/067,764	
	1998,4,29	09/069,308	
	1998,4,30	09/070,300	

有關微生物已寄存於： 寄存日期： 寄存號碼：

(請先閱讀背面之注意事項再填寫本頁各欄)

裝 訂 線

經濟部智慧財產局員工消費合作社印製

## 五、發明說明(1)

### 本發明之背景和概要

本發明一般係論及一種語音處理程序。詳言之，本發明係論及一種可用以產生彼等拼字之語音的系統。本發明可被採用在各種包括語音辨識、語音合成、和詞典編寫等多種不同之上下文意中。

在語音處理之業界中，彼等伴隨發音之拼字，係出現在許多不同之上下文意中。在語音辨識中，其字典內每一字詞有關之音標拼字，在使用前有必要被用來訓練其辨識器。傳統之音標拼字，係由彼等熟悉相關特定語言之音標拼字之絲毫差異的詞典編寫者親手建立。一字典要開發其每一字有關之良好音標拼字，將會是十分費時，以及需要大量之技術。若有一可靠之系統，可基於彼等之拼字來產生彼等字詞之音標拼字，則該等人力和特殊技能，大部份將可得以免除。該種系統可擴展彼等當前之辨識系統，使其可辨識現有字典內目前不能找到之地理位置和姓氏等字詞。

彼等拼字亦會經常發生在語音合成之業界中。今日之語音合成器，係藉著自一字典取出數位取樣之聲音單元，以及將此等聲音單元鏈結形成語句子，而將本文轉換成語音。

誠如以上諸例所展現，上述語音處理之語音辨識和語音合成，均可因其能自拼字產生精確發音之能力而得以獲利。然而，此一技術之需求則並非僅限於語音處理。詞典編寫者今日業已完成了許多主要世界語言相當大量及精確

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明（2）

之發音字典。然而，其中仍有數以百計之地方語言，缺乏良好之音標拼字。由於其生產一良好音標拼字之任務，迄今大部份仍屬人工做成，某些地方語言之轉譯，縱使有亦可能須要歷經數年方得以完成。若能有一良好之電腦具現技術，來評定轉譯之精確性，其轉譯程序將可得到大幅加速。此種評定系統，可使用一現有之語言轉譯資料集，來辨識其轉譯原型中發音有疑之字項。此將可大幅提昇一高品質轉譯成品之產生速率。

迄今在拼字發音之轉譯工作中的大部份嘗試，係單仰賴於彼等字母本身。此等技術離所希望者尚有一大段之距離。舉例而言，一唯文字發音產生器，要正確地發出字詞 Bible 之音，將會大有困難。僅基於字母之序列，上述之唯文字系統，將很可能會發出 "Bib-l" 之音，很像一學習認字之學齡兒童有可能發出之音。彼等傳統式系統中之難題，在於天生加諸在許多語言之發音規則上面之模糊性。以英語為例，其具有數以百計之發音規則，使得其逐字處理問題上自有其困難，以及在計算上所費不貲。

本發明係以一不同之角度來針對此一問題。本發明係使用一特別構成之混合決定樹，其涵蓋了文字序列和音素序列之決定下達法則。詳言之，此種混合決定樹在其樹之內部節點處，蘊含有一系列是-否之疑問。某些此等疑問涉及一拼字序列內之文字和彼等相鄰文字；其他疑問則涉及上述字序列中之音素和彼等相鄰音素；彼等內部節點最終將行至彼等之葉節點處，其將包含一所給文字，在彼等

### 五、發明說明(3)

文字序列所界定之字詞的發音上，最有可能之正確發音有關的概率。

本發明之發音產生器，係使用其混合決定樹來評定不同之發音候選者，而容許其能就一所給拼字選擇其最有可能之候選者，做為其最佳之發音。此最佳發音之產生，最好係屬一兩階段之程序，其中之第一階段，係使用一唯文字樹，來產生多數之發音候選者。此等候選者復在第二階段中，使用上述之混合決定樹，來選擇其最佳之候選者。

雖然其混合決定樹，係有利地使用在一兩階發音產生器中，其混合樹在解決某些並不需要唯文字之第一階段處理程序的問題中，係屬有用。舉例而言，其混合決定樹，可用以評定語言學者使用人工技術所產生之發音。

為更完全瞭解本發明、其諸目地和優點，可參照下列之說明和所附諸圖。

#### 圖示之簡要說明

第1圖係一可例示本發明之組件和步驟的方塊圖；

第2圖係一可例示一唯文字樹之樹狀圖；

第3圖係一可例示一依本發明所成之混合樹的樹狀圖

；

第4圖係一可例示一可用以產生上述依本發明所成之混合樹的當前較佳系統的方塊圖；

第5圖係一可例示一可用以經由一對齊程序而產生訓練資料的流程圖；

第6圖係一可例示其決定樹在一範例性發音產生器中

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明（4）

之用法之方塊圖；

第7圖係例示 Gini 準則在評估使用何者疑問聚居在一節點中之應用例；

第8圖係一依本發明所製文字聲音之發音產生器的方塊圖；而

第9圖則係一可例示一文字-語法-上下文意-方言之混合決定樹的樹狀圖。

較佳實施例之詳細說明

為例示本發明之原理，第1圖之範例性實施例，係顯示一拼字發音產生器。本發明之混合決定樹，誠如下文更詳細之解釋，可用於除本說明書所例示之發音產生器外之多種不同應用例中。此發音產生器被選來做例示說明，係由於其可強調出其混合決定樹結構之特徵和利益故也。

該發音產生器係採用兩個階段，其第一階段係採用一組唯文字決定樹10，以及其第二階段係採用一組混合決定樹12。一類似文字 B-I-B-L-E 序列之輸入序列14，係饋至一動態程式規劃音素序列產生器16。該發音產生器，係使用其唯文字決定樹10，來產生一表列發音18，其係表示上述拼字輸入序列之可能發音候選者。

該序列產生器，可依序檢查上述序列之每一文字，應用該文字相關之決定樹，基於其唯文字樹內所含有之概率資料，來選擇該文字所需之一音素發音。

上組唯文字決定樹，最好能包含每一字母文字有關之決定樹。第2圖係顯示文字 E 有關之唯文字決定樹的一個

## 五、發明說明 ( 5 )

範例。此決定樹係包含多數之內部節點 ( 圖中例示為橢圓形 ) 和多數之葉節點 ( 圖中例示為矩形 ) 。每一內部節點聚居有一是-否之疑問。在上述之唯文字樹中，此等疑問係針對所給之文字 ( 此一情況中之E文字 ) ，和彼等在上述輸入序列內之相鄰文字。注意在第2圖中，每一內部節點係依據相關疑問為是或否，而向左或向右分支。

第2圖中所用之縮寫符號如下：疑問中之數字，諸如 "+1" 或 "-1" ，係指稱相對於當前文字之拼字位置。舉例而言， "+1L== 'R' ?" 意謂『當前文字 ( 此一情況中之文字E ) 後之文字是否為一R』。彼等縮寫符號CONS和VOW，係表示彼等文字之類別，亦即，子音和母音。一相鄰文字之缺乏，或零文字，係以符號 "-" 表示，其係用做一可使某些文字與彼等對應音素發音對齊之填充字或佔位文字。符號 "#" 係表示一字詞之邊界。

彼等之葉節點，係聚居有概率資料，彼等可使彼等可能之音素發音，與彼等代表特定音素表示所給文字之正確發音的概率的數值相聯結。舉例而言，記號 "iy=>0.51" 意謂『此葉節點內之音素 'iy' 的概率係0.51』。其零音素，亦即靜音，係以 '-' 來表示。

其序列產生器16 ( 第1圖 ) ，因而係使用其唯文字決定樹10，來構成其表列18內所儲存之一個或多個發音臆測。每一發音最好能聯結一使用其決定樹10所選定各個音素之概率點數所達成之數值點數。彼等之字詞發音，可藉建立一可能組合之矩陣，以及接著使用動態程式規劃，而選擇

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明(6)

其n個最佳候選者，來加以計點。或者，彼等之n個最佳候選者，可使用一可先辨識其最有可能之字詞候選者，以及接著經由反覆取代，而產生額外之候選者的取代技術，來加以選擇。

藉著乘以其最高計點之音素（檢查彼等之葉節點辨得）的對應點數，首先選出其具有最高概率點數之發音，以及接著使用此一選出者，做為其最有可能之候選者，或第一最佳字詞候選者。藉著再次檢查彼等之葉節點內之音素資料，以辨識其先前未被選出，而與一最初選出之音素差異最小之音素，而接著選出彼等之額外（n個最佳之）候選者。此一最小差異之音素，接著取代上述之最初選出者，藉以產生其第二最佳字詞候選者。上述之程序可反覆重複，直至所需數目之n個最佳候選者被選出為止。其表列18可依漸小點數之順序來排序，以便其唯文字分析程序可判斷出之最佳發音，將會在該表列內首先出現。

誠如上文可注意到的，一唯文字分析程序，經常會產生不良之結果。此係由於上述之唯文字分析程序，在各文字處，無法決定因彼等後繼之文字將會產生之音素故也。因此，一唯文字分析程序，將會產生一自然語言中實際上不會發生之音素。舉例而言，一普通之名字Achilles，很有可能產生一標出ll之音的發音：ah-k-ih-ll-iy-z。在自然語言中，其第二個l實際上是不發音的：ah-k-ih-l-iy-z。上述使用唯文字樹之序列產生器，並不具有可篩選掉自然語言中決不會發生之字詞發音的機構。

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

線



## 五、發明說明 ( 7 )

上述發音系統之第二階段，係針對上述之問題。一混合樹點數預估器20，係使用上組混合決定樹12，來評估表列18內之每一發音的生命力。上述之點數預估器在工作上，係依序檢查上述輸入序列之每一文字，外加其序列產生器16指定給每一文字之音素。

如同上組之唯文字樹，上組混合樹係具有每一文字字母有關之混合樹。第3圖顯示一範例性混合樹。如同上述之唯文字樹，此混合樹具有一些內部節點和葉節點。在第3圖中，彼等之內部節點係例示成橢圓形，以及彼等葉節點係例示成矩形。彼等內部節點各係聚居有一是-否之疑問，以及彼等葉節點各係聚居有一些概率資料。雖然此混合樹之樹結構，係與上述唯文字樹者相類似，其中有一重要之差異。此混合樹之內部節點，可包含兩不同類別之疑問。一內部節點可包含一所給文字和其序列中相鄰文字有關之疑問，或者，其可包含該文字相關聯之音素，和彼等與其序列相對應之相鄰音素。其決定樹因而係混合成的，其將包含混合類別之疑問。

第3圖中所用之縮寫符號，係與第2圖中所用者相類似。其符號L係表示一文字和其相鄰文字有關之疑問。其符號P係表示一音素和其相鄰音素有關之疑問。舉例而言，其疑問"+1L=='D'?"意謂『在+1位置處之文字是否為一'D'?』。彼等縮寫符號CONS和SYL，係彼等音素之類別，亦即，子音和音節。舉例而言，"+1P==CONS?"意謂『在+1位置處之音素是否為一子音?』。彼等葉節點中之數字，係如

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明(8)

同彼等在唯文字樹中一樣，表示彼等音素之概率。

上述之混合樹點數預估器，可基於彼等混合樹之疑問，以及使用彼等混合樹之葉節點中的概率資料，重新計點其表列18內之每一發音。若有需要，彼等表列之發音，可如同表列22一樣，聯結其對應之點數一起被儲存。若有需要，該表列22可依漸小點數之順序來排序，以便其第一之表列發音，係其具有最高之點數者。

在許多案例中，其在表列22中含有最高點數位置之發音，將會不同於其在表列18中含有最高點數位置之發音。其發生係由於上述之混合樹點數預估器，使用其混合樹12，篩選掉該等未含有自相一致之音素序列的發音，或者彼等代表自然語言中決不會發生之發音故也。

若有需要，一選擇器模組24，可自由使用其表列22，以便自該表列，取出一個或多個發音。通常，該選擇器24係取出其具有最高點數之發音，以及將其提供做為輸出發音26。

誠如前文所應注意的，第1圖中所描繪之發音產生器，僅代表採用本發明之混合樹的一種可能實施例。以另一種實施例而言，其動態程式規劃音素序列產生器16，和其相關聯之唯文字決定樹10，在彼等一所給拼字序列早已有一個或多個現成之發音可用的應用例中，將可加以免除。此種情況係發生在現有一先前開發出之發音字典時。在此種情況下，其混合樹點數預估器20，和其混合決定樹12，可用來評定上述發音字典內之字項，辨識彼等具有低點數

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明(9)

者，藉以對所建字典內有疑問之發音加上旗標。此一系統舉例而言，可合併進一詞典編寫者之製造工具內。

上述之輸出發音或彼等選自表列22之發音，可用以形成語音辨識和語音合成兩者應用例有關之發音字典。在語音辨識之上下文意中，在其辨識器訓練片語期間，可使用上述之發音字典，來提供其辨識器字彙中未曾找到之字詞有關之發音。在語音合成之上下文意中，該等發音字典可用來產生彼等鏈結播放所需之音素聲音。此系統舉例而言，可用來增強一E-mail讀取器或其他本文轉換語音應用例之特徵。

本發明之混合樹評定系統，可用於希望單一或表列之可能發音的多種應用例中。舉例而言，在一動態線上字典中，其使用者可鍵入一字詞，以及該系統可提供一表列按概率順序排列之可能發音。該評定系統，亦可用做彼等語言學習系統所需之一使用者回授工具。一具有語音辨識能力之語言學習系統，係用來顯示一拼字，以及用來分析一說話者在發上述新語言中之該字詞的音時的企圖，以及該系統將會告訴該使用者，其對該字詞之發音係有多大之可能性或多大之不可能性。

### 彼等決定樹之產生

第4圖顯示上述可產生唯文字樹和混合樹之系統。在此決定樹產生系統之中央，係樹產生器40。此樹產生器係採用一樹成長演算法，其可在該系統開發者所提供預定組之訓練資料42上面運作。通常，該訓練資料係包含對齊之

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明 ( 10 )

文字、音素對，彼等係對應於字詞已知之適當發音。該訓練資料可經由第5圖中所例示之對齊程序來產生。第5圖例示一正在一範例性字詞BIBLE上面執行之對齊程序。其拼字44和其發音46，係饋至一動態程式規劃對齊模組48，其可使該拼字之文字與其對應發音之音素對齊。理應注意的是，在所例示之範例中，其最後之E係不發音。該等文字音素對，接著便會儲存做為資料42。

回顧第4圖，該樹產生器在工作上係配合三個額外之組件：一組可能之是-否疑問50、一組可就每一文字選擇最佳疑問或可決定該節點是-否應為一葉節點之規則52、和一可防止訓練過度之刪除法53。

該組可能之是-否疑問，依據正在成長的，係一唯文字樹，或一混合樹而定，可包含彼等文字疑問54和音素疑問56。當成長一唯文字樹時，將僅有文字疑問54被使用；當成長一混合樹時，彼等文字疑問54和音素疑問56，將同被使用。

在此一當前較佳實施例中，彼等選擇其聚居在每一節點處之最佳疑問，在設計上係遵從上述之Gini準則。其他之分割準則可代而用之。進一步有關分割準則之資訊，可參考Breiman, Friedman et al, 之 "Classification and Regression Trees"。基本上，上述之Gini準則，係用來自上組可能之是-否疑問50選擇一疑問，以及採用一決定何時一節點係一葉節點之停止規則。上述之Gini準則，係採用一稱做『雜度』之觀念。此一雜度永遠係一非負數之數

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明 ( 11 )

目字。其應用在一節點上，係使一包括等比例之所有可能範疇之節點，係具有最大之雜度，而一僅包括彼等可能範疇中之一範疇的節點，係具有零值之雜度（最小之可能值）。在此有數種函數可滿足上述之條件。彼等係依一節點內每一範疇之計數而定。上述之Gini準則，可界定如下。若C係資料項目可隸屬之類別組，以及T係當前之樹節點，令 $f(1|T)$ 為在T節點中隸屬類別1之訓練資料項目， $f(2|T)$ 為在T節點中隸屬類別2之訓練資料項目，等等，則

$$i(T) = \sum_{j,k \in C, j \neq k} f(j|T)f(k|T) = 1 \sum_j [f(j|T)]^2$$

藉範例來例示說明，假定該系統係就"E"成長一樹。在該樹之一所給之節點T中，該系統舉例而言，可能具有"E"在該字詞中如何發音之10個範例。舉例而言，在此等範例之5中，"E"係發"iy"（"cheeze"中之"ee"音）之音；在此等範例之3中，"E"係發"eh"（"bed"中之"e"音）之音；以及在此等範例之中，"E"係發"-"（如同"maple"中"e"之不發音）。

假定該系統正在考慮兩個可應用在上述10個範例中之可能的是-否疑問， $Q_1$ 和 $Q_2$ 。彼等對 $Q_1$ 回答『是』之項目，將包含四個"iy"之範例，和一個"-"之範例（其他五個項目對 $Q_1$ 回答『否』）。彼等對 $Q_2$ 回答『是』之項目，將包含三個"iy"之範例，和三個"eh"之範例（其他四個項目對 $Q_1$ 回答『否』）。第6圖係以示意圖比較此兩種情況。

上述之Gini準則，可回答該系統應就此一節點選擇何者疑問， $Q_1$ 或 $Q_2$ 。上述用以選擇正確疑問之Gini準則是：

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

線

## 五、發明說明 ( 12 )

找出自父節點行至子節點其雜度下降最大之疑問。此一雜度下降  $\Delta I$  係界定為  $\Delta I = i(T) - P_{yes} * i(yes) - P_{no} * i(no)$ ，此處， $P_{yes}$  係彼等行至「是」子節點之項目的比例值，以及  $P_{no}$  係彼等行至「否」子節點之項目的比例值。

應用 Gini 準則至上述之範例：

$$i(T) = 1 - \sum_j [f(j/T)]^2 = 1 - 0.5^2 - 0.3^2 - 0.2^2 = 0.62$$

就  $Q_1$  而言， $\Delta I$  因而成為：

$$i(T) - P_{yes}(Q_1) = 1 - 0.8^2 - 0.2^2 = 0.32$$

$$i(T) - P_{no}(Q_1) = 1 - 0.2^2 - 0.6^2 = 0.56$$

$$\text{故 } \Delta I(Q_1) = 0.62 - 0.5 * 0.32 - 0.5 * 0.56 = 0.18。$$

$$\text{就 } Q_2 \text{ 而言，吾等可得 } \Delta I(\text{yes}, Q_2) = 1 - 0.5^2 - 0.5^2 = 0.5，$$

$$\text{以及 } i(\text{no}, Q_2) = (\text{同式}) = 0.5。$$

$$\text{故 } \Delta I(Q_2) = 0.62 - (0.6) * (0.5) - (0.4) * (0.5) = 0.12。$$

在此一情況下， $Q_1$  在雜度上將會產生最大之下降。

所以，其將會被選擇而非  $Q_2$ 。

上述之規則組 52 將可就一節點，聲明其自父節點行至子節點而招致雜度下降最大之疑問，為一最佳之疑問。

上述之樹產生器，可應用彼等之規則 52，以成長一選自上組 50 之是-否疑問的決定樹。該產生器將會使該樹繼續成長，直至其最佳尺寸之樹已成長為止。彼等之規則 52 包含一組停止規則，彼等可在該樹已成長至一預定尺寸時，使該樹停止成長。在此一較佳之實施例中，該樹可成長至一大於最終所希望之尺寸。接著，彼等修剪法 53，可用

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

### 五、發明說明 ( 13 )

以將該樹修剪回其所希望之尺寸。上述之修剪法，可具現上述所引用之參考文獻中所描述之Breiman技術。

上述之樹產生器，因而可依據上組可能之是-否疑問50，係單獨包含唯文字疑問，亦或包含結合音素疑問之唯文字疑問，來產生60處一般所示之數組唯文字樹，或70處一般所示之混合樹。其訓練資料42之資料集，係如上文所述，包含有文字、音素對。在一唯文字樹之成長中，僅有彼等成對之文字部分，被用來聚居彼等之內部節點。反之，當成長彼等之混合樹時，彼等訓練資料對之文字和音素兩成分，均可能被用來聚居彼等之內部節點。在兩者情況中，彼等成對之音素部分，係被用來聚居彼等之葉節點。彼等與葉節點內之音素資料相關聯之概率資料，係藉計數在整個訓練資料之資料集中，一所給音素與一所給文字發生對齊之次數，來加以產生。

上述方法所產生之文字發音決定樹，可儲存在記憶體內，以供各種不同語音處理應用例使用。雖然此等應用例為數眾多及多樣化，下文將出示數例，以凸顯此等樹之某些能力和優點。

第6圖係例示彼等唯文字樹和混合樹，兩者均被用來自拼字文字序列產生出彼等之發音。雖然所例示之實施例，係對唯文字樹和混合樹兩者一起加以採用，其他之應用例則可能僅使用到一成分，而用不到另一成分。在所例示之實施例中，其唯文字樹組係儲存在80處之記憶體內，及其混合樹係儲存在82處之記憶體內。在許多應用例中，

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明 ( 14 )

其每一字母文字將會有一樹。其動態程式規劃序列產生器84，可對輸入序列86運作，以基於其唯文字樹80，而產生一88處之發音。基本上，上述輸入序列內之每一文字，係各別做考慮，以及上述之唯文字樹，係用來選擇該文字最有可能之發音。誠如前文所述，其唯文字樹將會就上述序列內之所給文字和其相鄰文字，尋問一系列之是-否疑問。在該序列內之所有文字均考慮過後，其最後之發音將可藉鏈結上述序列產生器所選定之音素而產生出。

欲改進其發音，可使用上述之混合樹組82。其唯文字樹雖僅能尋問有關文字之疑問，其混合樹則可尋問有關文字及有關音素之疑問。其計點器90可自其序列產生器84之輸出，接收到音素資訊。就此點而言，其序列產生器84，使用其唯文字樹80，可產生多數之不同發音，而基於彼等各自之概率點數，來排序該等發音。此一排序過之發音表列，可儲存在92處，以供其計點器90自由使用。

其計點器90可接收上述同時供應至序列產生器84之輸入序列86，而做為其輸入。其評定器90可將彼等混合樹82之疑問，應用至上述之文字序列，而在被尋及嚮應一音素疑問時，使用上述來自儲存器92之資料。其94處之最後輸出，通常，係一較88處所供應者為佳之發音。其理由在於彼等之混合樹易於篩選掉自然語言中不會發生之發音。舉例而言，上述之普通名字，Achilles，將很有可能產生一標出ll之音的發音：ah-k-ih-ll-iy-z。在自然語言中，其第二個l實際上是不發音的：ah-k-ih-l-iy-z。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線



## 五、發明說明 ( 15 )

若有需要，其點數產生器90，亦可製造一如同在96處之一排序過之n個可能發音的表列。彼等與各發音相關聯之點數，係代表該發音內指定給每一音素之個別概率點數的合成值。此等點數本身，可被用在有疑問之發音需要被辨識的應用例中。舉例而言，一詞典編寫團隊所提供之音標拼字，可使用該等混合樹迅速辨識任何有疑問之發音，來加以檢核。

### 文字-聲音發音產生器

為例示本發明之原理，第8圖之範例性實施例，顯示有一兩階段拼字文字-發音產生器。誠如下文更詳細之解釋，本發明之混合決定樹的處理方法，可用於除本說明書所例示之發音產生器外之多種不同之應用例中。此一兩階段發音產生器被選來做例示說明，係由於其可強調出其混合決定樹結構之特徵和利益故也。

上述之兩階段發音產生器包含：一第一階段116，其最好係採用一組文字-語法-上下文意-方言之決定樹110；和一第二階段120，其係採用一組可在一音素水平下檢查其輸入序列114之音素混合決定樹112。彼等文字-語法-上下文意-方言之決定樹，可檢查一拼字序列中彼等涉及文字和彼等相鄰文字之疑問（亦即，文字相關之疑問）；其他受到檢查之疑問有，一特定字詞前接或後接者為何字詞（亦即，上下文意相關之疑問）；再一其他受到檢查之疑問有，該字詞在一句子內為語言之何一部分，加上其他字詞在該句子內具有何種語法（亦即，語法相關之疑問）；

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明 ( 16 )

再一其他受到檢查之疑問有，其所希望說出為何種方言。  
一使用者最好藉方言選擇裝置150，來選擇所要說出為何種方言。

本發明之另一實施例包含：使用彼等文字相關之疑問，和至少一字詞水平特性（亦即，語法相關之疑問，或上下文意相關之疑問）。舉例而言，一實施例在其第一階段係利用一組文字-語法決定樹。另一實施例係利用一組不檢查其輸入序列之語法的文字-上下文意-方言決定樹。

理應瞭解的是，本發明並非僅限於彼等發生在一句子中之字詞，而係包括其他呈現語法之語言結構，諸如片段之句子或片語等。

一輸入序列114，諸如一句子之文字序列，係饋送至其基於本文之發音產生器116。舉例而言，該輸入序列114，可能會是下列之句子："Did you know who read the autobiography?"。

一語法資料115，係其基於本文之發音產生器116之輸入。此一輸入可提供其基於本文之發音產生器116有關之資訊，藉以正確地循經其文字-語法-上下文意-方言之決定樹110。該語法資料115，可宣告其輸入序列114中之每一字詞，具有語音之何一部分。舉例而言，上述輸入序列範例中之字詞"read"，將會被其語法標記器軟體模組129，標記為一動詞（而非一名詞或一形容詞）。上述之語法標記器軟體技術，可得自"Xtag"計劃下類似賓州大學等之機構。此外，以下之參考資料亦討論到語法標記器軟體技

（請先閱讀背面之注意事項再填寫本頁）

裝

訂

線

## 五、發明說明 ( 17 )

術：1991年11月11日加拿大蒙特利爾McGill大學所出版George Foster電腦科學之碩士論文"Statistical Lexical Disambiguation"。

其基於本文之發音產生器116，係使用上述之決定樹110，來產生一表列發音118，其係代表上述拼字輸入序列之可能發音候選者。該表列118之每一發音（例如，發音A），係表示上述最好能包含每一字詞如何標出重音之輸入序列114的發音。此外，此一實施例可決定出每一字詞說出之速率。

其句子速率計算器軟體模組152，係被其基於本文之發音產生器116，用來決定每一字詞應說出多快。舉例而言，其句子速率計算器152，可檢查句子之上下文意，以決定該句子中之某些字詞，應以較正常者說出的較快些或較慢些。舉例而言，一在尾部具有一感嘆符號之句子，其將會提出該句子結尾前一預定數目之字詞，當較正常者具有較短之期間的速率資料，以便表達出一感嘆陳述之影響力。

其基於本文之發音產生器116，可依次檢查上述序列之每一文字和字詞，應用其與該文字或字詞之語法（或字詞之上下文意）相關聯之決定樹，基於該混合樹內所包含之概率資料，來選擇該文字有關之一音素發音。該組決定樹110，最好包含所涉及語言之字母和語法內之每一文字有關的決定樹。

第9圖顯示一可應用至上述字詞"READ"中之文字"E"

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明 ( 18 )

的範例性文字-語法-上下文意-方言之決定樹140。此混合樹包含多數之內部節點 (圖中以橢圓形例示) 和多數之葉節點 (圖中以矩形例示)。各內部節點聚居有一是-否之疑問。此等是-否之疑問，係一些可回答是或否之疑問。在該文字-語法-上下文意-方言之決定樹140中，此等疑問係針對：上述輸入序列中之一所給文字 (例如，此一情況下中之文字"E") 和其相鄰之文字；或該句子中字詞之語法 (例如，名詞、動詞、等等)；或該句子之方言。在第9圖中理應注意的是，各內部節點依據對其相關聯疑問之回答係是亦否而定，將會向左或向右分支。

其第一內部節點最好訊問要說出之方言。其內部節點138係此種訊問之代表。若要說出的是南部方言，則其將循經一南部方言混合樹139，其最終在彼等葉節點處，將會呈現一更具南部方言之特色的音素值。

第9圖中所用之縮寫符號如下：疑問中之數字，諸如"+1"或"-1"，係指稱相對於當前文字之拼字位置。符號L係表示一文字和其相鄰文字有關之疑問。舉例而言，"-1 == 'R' or 'L'?" 意謂『當前文字 (其為'E') 前之文字係'L'亦或'R'』。彼等縮寫符號CONS和VOW，係表示彼等文字之類別，亦即，子音和母音。符號"# "係表示一字詞之邊界。術語'tag(i)' 係表示一有關其第i個字詞之語法標記的疑問，其中，i=0係表示當前之字詞，i=-1係表示其前接之字詞，i=+1係表示其後接之字詞，等等。因此，"tag(0) == PRES?" 係表示『當前之字詞是否為一現在式動詞?』。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明 ( 19 )

彼等葉節點聚居有一些可使彼等可能之音素發音，與彼等表示該特定音素代表所給文字之正確發音之概率的數字值相結合。其零音素，亦即不發音，係以符號“-”表示。

舉例而言，現在式動詞"READ"和"LEAD"中之"E"，被其決定樹140在葉節點142處，指定其具有概率1.0之正確發音"iy"，過去式動詞"READ"中之"E"，則會在葉節點144處，被指定具有概率0.9之發音"eh"。

其決定樹110 (第8圖者)，最好包含上下文意相關之疑問。舉例而言，彼等內部節點之上下文意相關疑問，可檢查字詞"you"是-否前接字詞"did"。在此種上下文意中"you"中之"y"，在俗語中通常係發成"ja"之音。

本發明亦可產生詩體性質之資料，藉以表達說出一句子時之重音、揚音、抑音、或停頓等特徵。此等語法相關之疑問，將有助於決定該音素如何表達出重音、或揚音或抑音。舉例而言，其內部節點141 (第9圖者)，將會訊問該句中之第一字是否為一疑問代名詞，諸如範例性句子"Who read a book?"中之"who"。由於在此一範例中，此一範例中之第一字為一疑問代名詞，故選擇其具有音素重音之葉節點144。其葉節點146係例示其他不加重音之音素的選擇項。

就另一範例而言，在一疑問句中，其最後一字之最後一音節的音素，該具有一揚音符號，以便更能表達出該句之疑問特徵。再一範例可使本發明包括能在說出一句子時

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明(20)

能容有一自然停頓。本發明係藉尋求有關標點符號，諸如逗點和句點之疑問，而納入此種停頓之細節。

其基於本文之發音產生器116(第8圖)，因而可使用決定樹110，來建立一個或多個發音之假設，而將其儲存進其表列118中。每一發音最好能與一藉著結合該等使用決定樹110所選擇之各個音素的概率點數所達成之數字點數相關聯。彼等字詞發音可藉建立一可能結合之矩陣，再使用動態程式規劃，選出n個最佳候選者，而加以計點。

或者，彼等n個最佳候選者在選擇上，可使用一取代技術，其將會在下列之方式中，首先辨識出最有可能之字詞候選者，以及接著經由反覆取代而產生額外之候選者。

藉著乘以其最高計點之音素(檢查彼等之葉節點辨識出的)的對應點數，首先選出其具有最高概率點數之發音，以及接著使用此一選出者，做為其最有可能之候選者，或第一最佳字詞候選者。藉著再次檢查彼等之葉節點內之音素資料，以辨識其先前未被選出而與一最初選出之音素差異最小之音素，而接著選出彼等之額外(n個最佳之)候選者。此一最小差異之音素，接著取代上述之最初選出者，藉以產生其第二最佳字詞候選者。上述之程序可反覆重複，直至所需數目之n個最佳候選者被選出為止。其表列118可依漸小點數之順序來排序，以便其唯文字分析程序可判斷出之最佳發音，將會在該表列內首先出現。

其決定樹110，經常僅能製造出中適成功之結果。此係由於此等混合樹無法在每一文字處，決定出彼等後繼之

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明(21)

文字將會產生何種音素故也。因此，其決定樹110，有可能產生一高計點之發音，其實際上卻不會發生在自然語言中。舉例而言，上述之普通名字Achilles，很有可能產生一標出ll之音：ah-k-ih-ll-iy-z。在自然語言中，其第二個l實際上是不發音的：ah-k-ih-l-iy-z。上述使用決定樹110之發音產生器，並不具有可篩選掉自然語言中決不會發生之字詞發音的機構。

上述發音系統108之第二階段120，係針對上述之問題。一音素混合樹點數預估器120，係使用上組音素混合決定樹112，來評估表列118內之每一發音的生命力。上述之點數預估器120在工作上，係依序檢查上述輸入序列114之每一文字，外加其基於本文之發音產生器116指定給每一文字之音素。

上述之音素混合樹點數預估器120，可基於彼等音素混合樹之疑問112，以及使用彼等混合樹之葉節點中的概率資料，重新評定其表列118內之每一發音。若有需要，彼等表列之發音，可如同表列122一樣，聯結其對應之點數一起做儲存。若有需要，該表列122可依漸小點數之順序來排序，以便其第一之表列發音，係其具有最高點數者。

在許多案例中，其在表列122中含有最高點數位置之發音，將會不同於其在表列118中含有最高點數位置之發音。其之發生係由於上述之音素混合樹點數預估器120，使用其音素混合樹112，篩選掉該等未含有自相一致之音

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明 ( 22 )

素序列的發音，或者彼等代表自然語言中決不會發生之發音故也。

在此一較佳實施例中，其音素混合樹點數預估器120，係依序使用其句子速率計算器152，來決定其表列122中之發音有關的資料。此外，該預估器120，係在類似於上述處理方法之方式下，來利用彼等可容有要檢查方言有關之疑問，和彼等亦可在葉節點處容有用以決定重音與其他詩體特徵之疑問的音素混合樹。

若有需要，一選擇器模組124可自由使用其表列122，以便自該表列取出一個或多個發音。通常，該選擇器124係取出其具有最高點數之發音，以及將其提供做為輸出發音126。

誠如前文所應注意的，第8圖中所描繪之發音產生器，僅代表採用本發明之混合樹處理方法的一種可能實施例。在另一種實施例中，上述之輸出發音或彼等選自表列22之發音，可用以形成語音辨識和語音合成兩者應用例有關之發音字典。在語音辨識之上下文意中，在其辨識器訓練片語期間，可使用上述之發音字典，來提供其辨識器字彙中未曾找到之字詞有關之發音。在語音合成之上下文意中，該等發音字典可用來產生彼等鏈結播放所需之音素聲音。此系統舉例而言，可用來增強一E-mail讀取器或其他本文轉換語音應用例之特徵。

本發明之混合樹評定系統（亦即，文字、語法、上下文意、和音素），可用於希望單一或表列之可能發音的多

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線



### 五、發明說明 ( 23 )

種應用例中。舉例而言，在一動態線上語言學習系統中，一使用者可鍵入一句子，以及該系統可提供一表列有關該句子按概率順序排列之可能發音。該評定系統，亦可用做彼等語言學習系統所需之一使用者回授工具。一具有語音辨識能力之語言學習系統，係用來顯示一拼字，以及用來分析說話者在發出上述新語言中該句子之音時的企圖。該系統將會指示上述之使用者，其對該句子之發音係有多大之可能或多大之不可能。

雖然本發明業已在其當前之較佳形式中做了說明，理應瞭解的是，其混合樹發音系統係有數目眾多之應用例。因此，本發明在不違離本發明之精神下，依據其所附之申請專利範圍，將可有某些之修飾體和變更形式。

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

## 五、發明說明(24)

## 元件編號對照

- |                        |                                 |
|------------------------|---------------------------------|
| 10...唯文字決定樹            | 92...儲存器                        |
| 12...混合決定樹             | 84...動態程式規劃序列產生器                |
| 14...輸入序列              | 86...輸入序列                       |
| 16...動態程式規劃音素序列<br>產生器 | 88...發音                         |
| 18,22...表列發音           | 90...計點器                        |
| 20...混合樹點數預估器          | 110,140...文字-語法-上下文意-<br>方言之決定樹 |
| 24...選擇器模組             | 112...音素混合決定樹                   |
| 26...輸出發音              | 114...輸入序列                      |
| 40...樹產生器              | 115...語法資料                      |
| 42...訓練資料              | 116...基於本文之發音產生器                |
| 44...拼字                | 118...表列發音                      |
| 46...發音                | 120...音素混合樹點數預估器                |
| 48...動態程式規劃校準模組        | 122...表列發音                      |
| 50...可能之是-否疑問          | 124...選擇器模組                     |
| 52...規則                | 126...輸出發音                      |
| 53...修剪法               | 129...語法標記器軟體模組                 |
| 54...文字疑問              | 138,141...內部節點                  |
| 56...音素疑問              | 139...南部方言混合樹                   |
| 60,80...唯文字樹           | 142,144,146...葉節點               |
| 70,82...混合樹            | 150...方言選擇裝置                    |
| 80,82...記憶體            | 152...句子速率計算器軟體模組               |

(請先閱讀背面之注意事項再填寫本頁)

裝

訂

線

四、中文發明摘要(發明之名稱： 利用決定樹來產生及評定拼字的多重發音的方法與裝置 )

本混合決定樹係包含一種有關一拼字序列內相鄰文字，和有關與此拼字序列相對應之音素序列內之相鄰音素的是-否疑問的網絡。此混合決定樹之葉節點，可提供何者音標拼字最具可能性之相關資訊。使用此等混合樹，每一多數可能之發音均可發展出一些點數，以及此等點數不但可用以選擇其最佳之發音，亦可用以使彼等發音按概率之高低而分出等級。此系統所產生之發音，不但可用在語音合成和語音辨識之應用例中，亦可用在詞典編寫之應用例中。

英文發明摘要(發明之名稱： METHOD AND APPARATUS USING DECISION TREES TO GENERATE AND SCORE MULTIPLE PRONUNCIATIONS FOR A SPELLED WORD )

The mixed decision tree includes a network of yes-no questions about adjacent letters in a spelled word sequence and also about adjacent phonemes in the phoneme sequence corresponding to the spelled word sequence. Leaf nodes of the mixed decision tree provide information about which phonetic transcriptions are most probable. Using the mixed trees, scores are developed for each of a plurality of possible pronunciations, and these scores can be used to select the best pronunciation as well as to rank pronunciations in order of probability. The pronunciations generated by the system can be used in speech synthesis and speech recognition applications as well as lexicography applications.

(請先閱讀背面之注意事項再填寫本頁各欄)

裝

訂

線

## 六、申請專利範圍

1. 一種可就自一預定字母所選出之文字輸入序列產生至少一發音之裝置，其包含：

一記憶體，其可用以儲存與該字母相對應之多數唯文字決定樹，

該等唯文字決定樹，係具有彼等可代表一所給序列中之所給文字和其相鄰文字有關之是-否疑問的內部節點；

該記憶體尚可儲存與該字母相對應之多數混合決定樹，

此等混合決定樹，係具有第一多數可代表上述所給序列中之所給文字和其相鄰文字有關之是-否疑問的內部節點，以及具有第二多數可代表上述所給序列中之所給音素和其相鄰音素有關之是-否疑問的內部節點；

彼等唯文字決定樹和混合決定樹，尚具有彼等可代表一些可使上述所給文字與多數音素發音相關聯之概率資料的葉節點；

一音素序列產生器，其係與該唯文字決定樹相耦合，可用以處理一文字輸入序列，以及可用以產生第一組與該文字輸入序列相對應之發音；和

一點數預估器，其係與該混合決定樹相耦合，可用以處理上述之第一組，以產生一第二組計點之發音，此等計點之發音，係代表該文字輸入序列至少之一發音。

(請先閱讀背面之注意事項再填寫本頁)

裝  
訂  
線

## 六、申請專利範圍

9. 如申請專利範圍第1項所申請之裝置，其中尚包含一語音辨識系統，其具有一可用以做其辨識器訓練用之發音字典，以及此字典至少聚居有上述第二組之一部分，以便提供彼等字詞基於彼等拼字之發音。
10. 如申請專利範圍第1項所申請之裝置，其中尚包含一語音合成系統，其可用以接收至少上述第二組之一部分，以便產生彼等字詞基於彼等拼字之聲音合成發音。
11. 如申請專利範圍第10項所申請之裝置，其中之語音合成系統，係合併在一E-mail讀取器內。
12. 如申請專利範圍第10項所申請之裝置，其中之語音合成系統，係合併在一字典內，以便提供一表列按概率順序排列之可能發音。
13. 如申請專利範圍第1項所申請之裝置，其中尚包含一語言學習系統，其可用以顯示一拼字，以及可使用該等唯文字決定樹和混合決定樹，來分析一說話者在發出上述字詞之音時的企圖，以便告訴該說話者，其對該字詞之發音有多大之可能性。
14. 一種製造拼字-發音資料之方法，其包含之步驟有：
  - 提供一第一組有關一輸入序列之文字和彼等與相鄰文字之關係的是-否疑問；
  - 提供一第二組有關一輸入序列之音素和彼等與相鄰音素之關係的是-否疑問；
  - 提供一可代表多數不同組對之訓練資料集，彼等每一對係包含一文字序列和一音素序列，該文字序列

(請先閱讀背面之注意事項再填寫本頁)

裝 · · · · · 訂 · · · · · 線

## 六、申請專利範圍

上述基於本文之決定樹，係具有一些代表彼等與上述輸入序列之預定特性有關之疑問的內部節點；

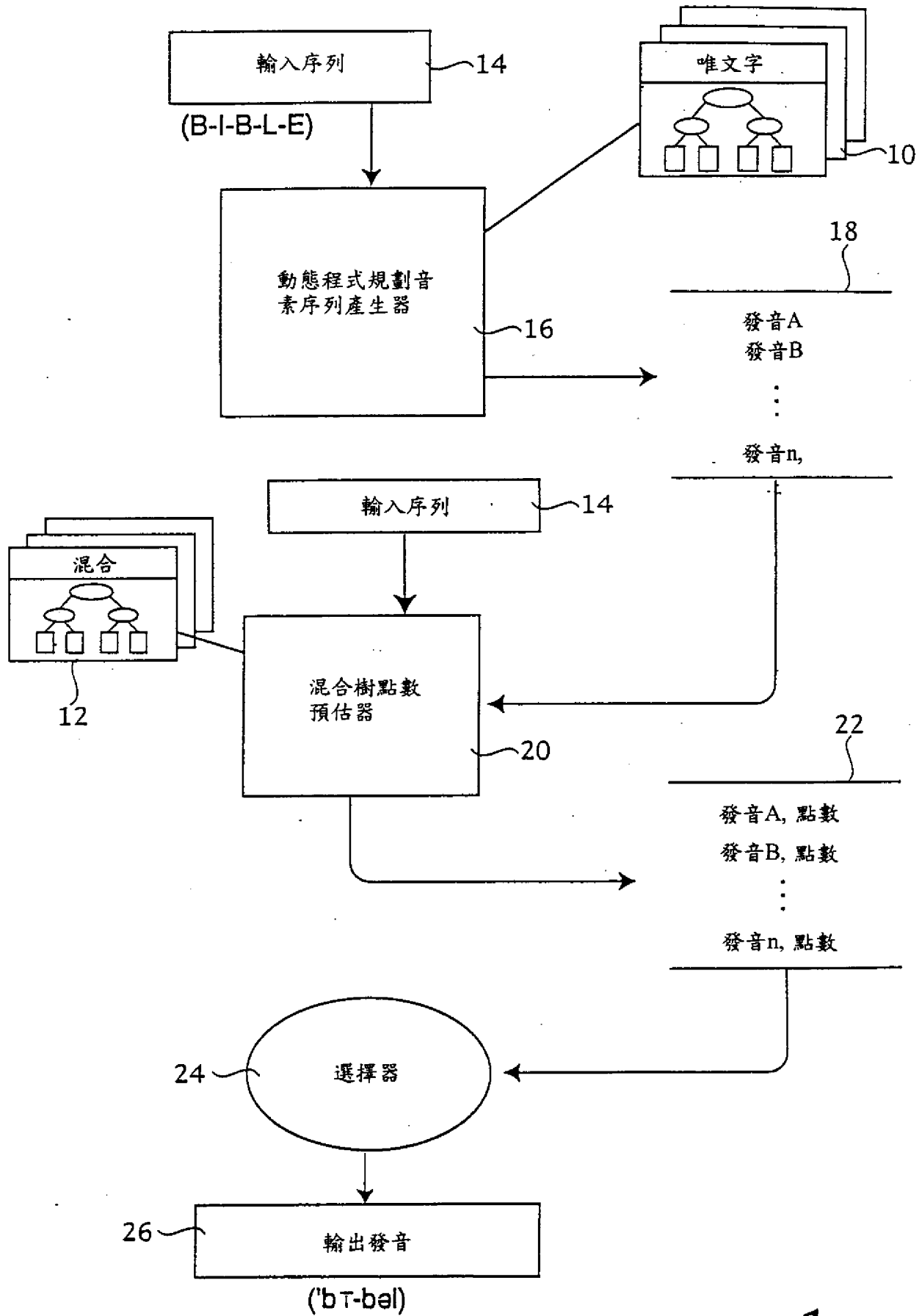
上述基於本文之決定樹，尚具有一些代表彼等可使每一文字與多數音素發音相關聯之概率資料的葉節點；和

一基於本文之發音產生器，其係與上述基於本文之決定樹相連接，可用以處理上述之文字輸入序列，以及可基於上述基於本文之決定樹，來產生一第一組與上述之文字輸入序列相對應之發音。

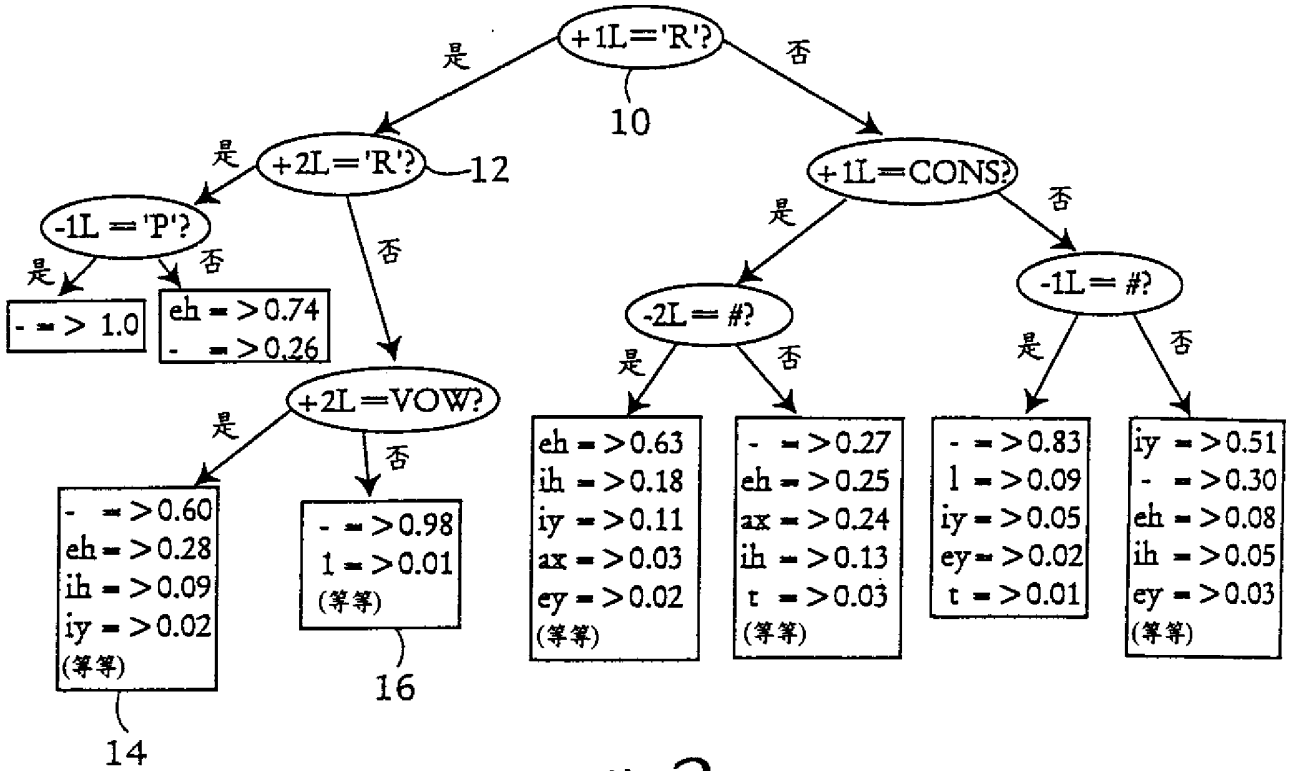
23. 如申請專利範圍第22項所申請之裝置，其中尚包含一音素混合樹點數預估器，(其係與上述基於本文之發音產生器相連接，)可用以處理上述之第一組，以產生一第二組計點之發音，此等計點之發音，係代表該文字輸入序列至少之一發音。

(請先閱讀背面之注意事項再填寫本頁)

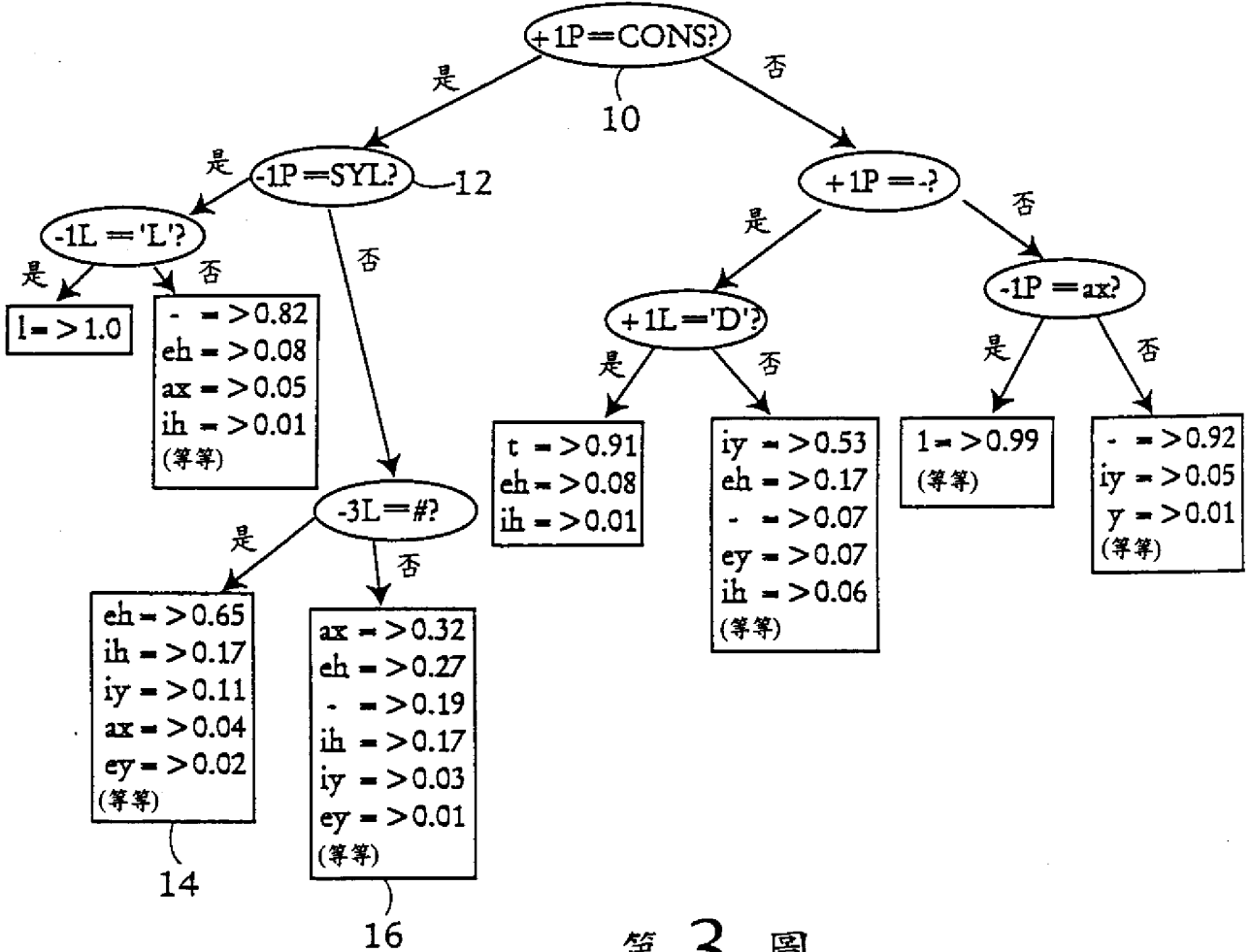
裝  
訂  
線



第 1 圖

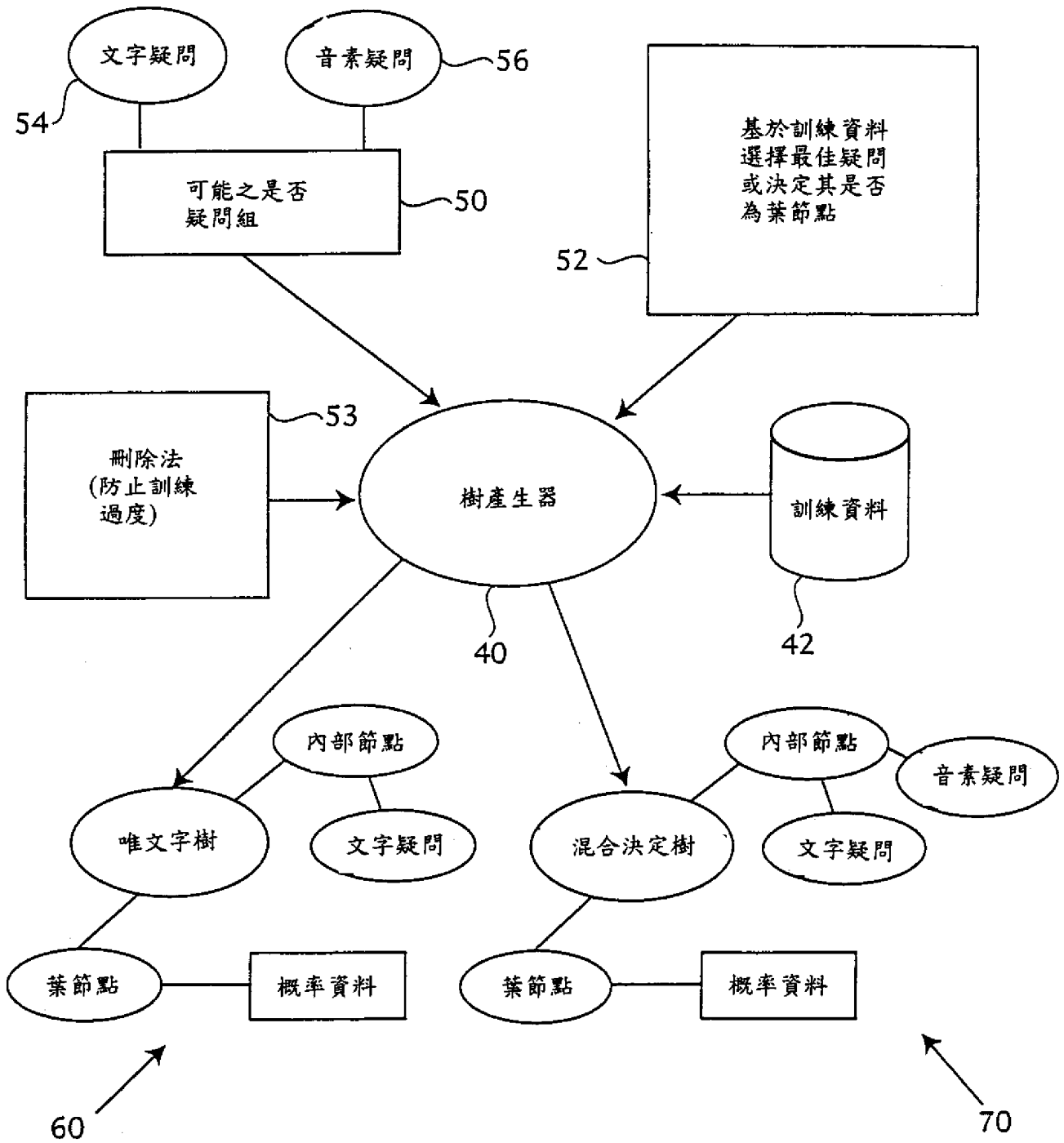


第 2 圖

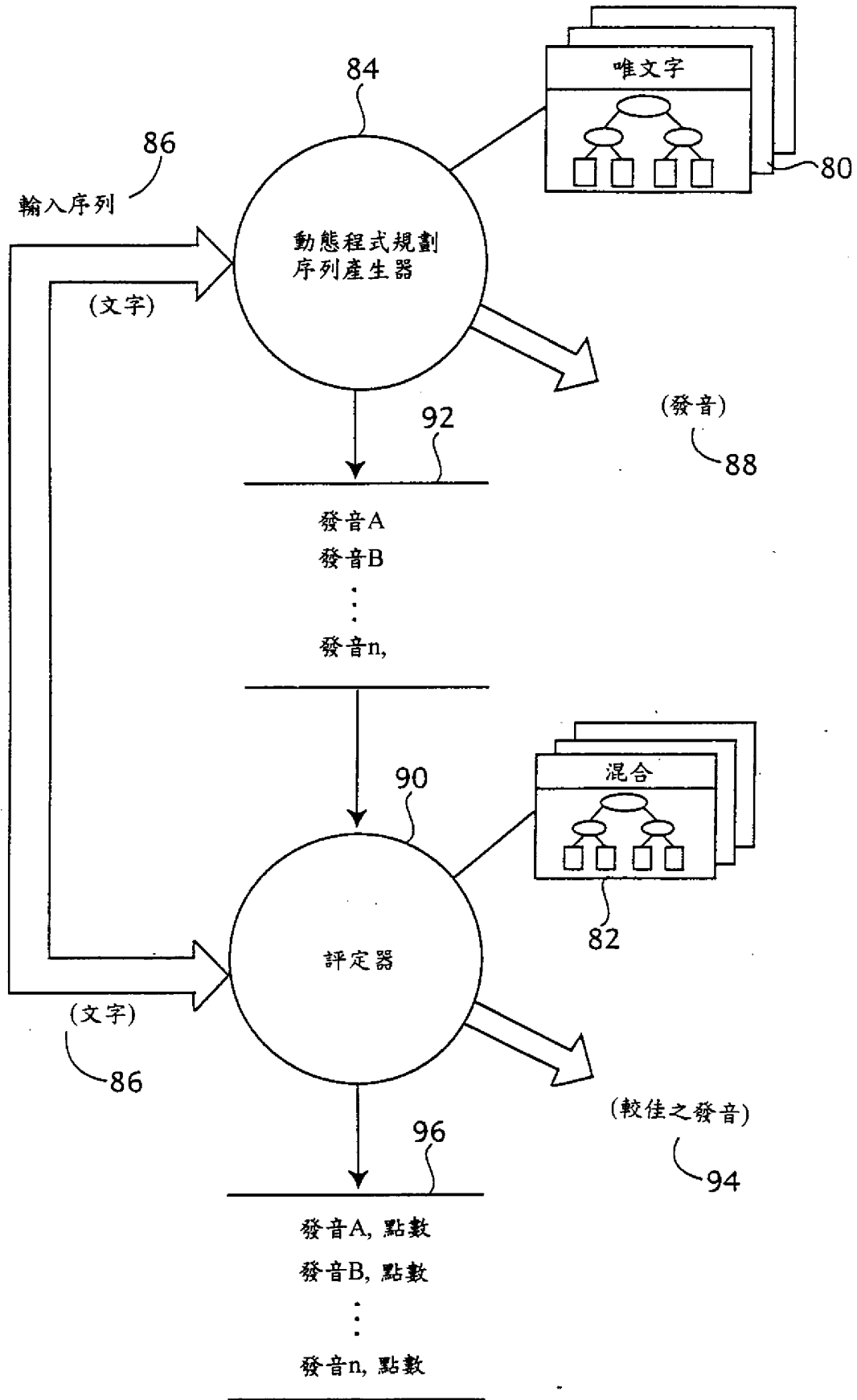


第 3 圖

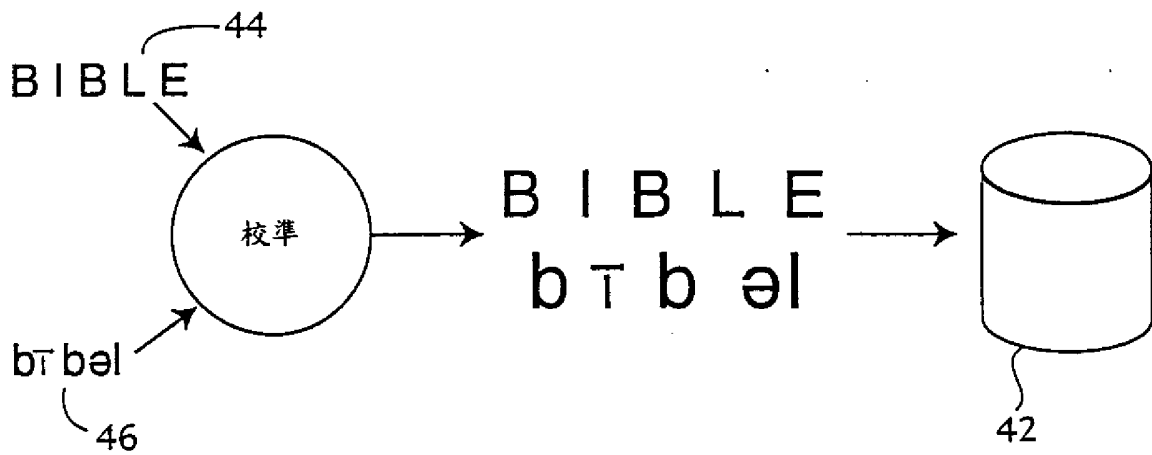




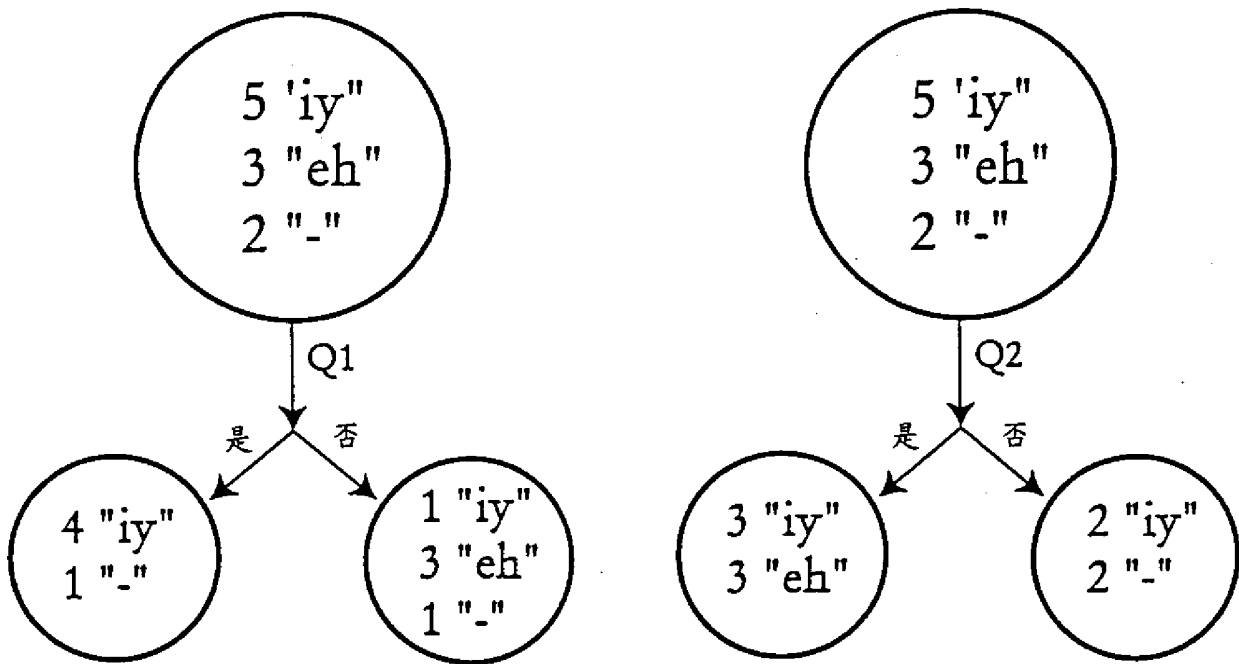
第 4 圖



第 6 圖

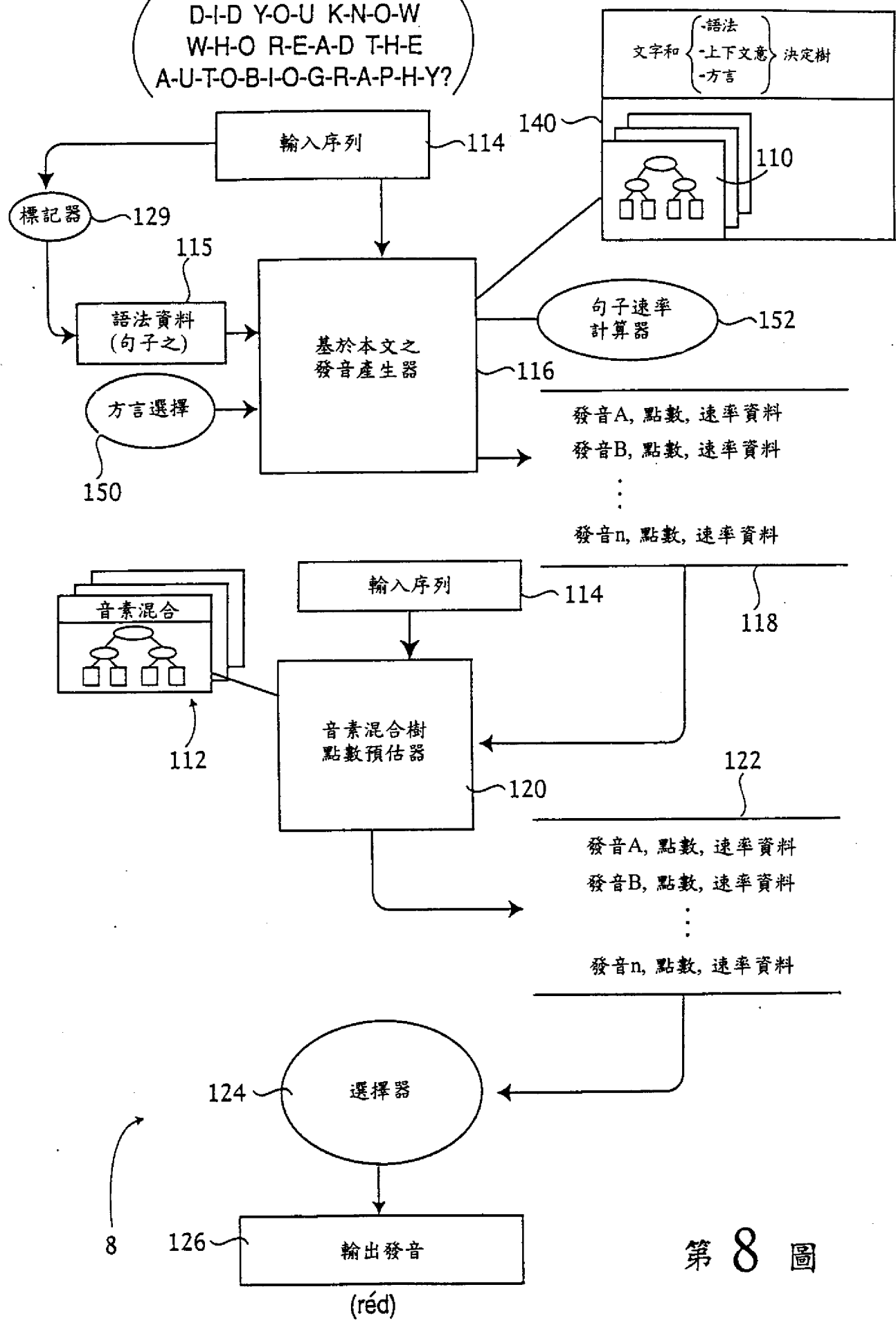


第 5 圖

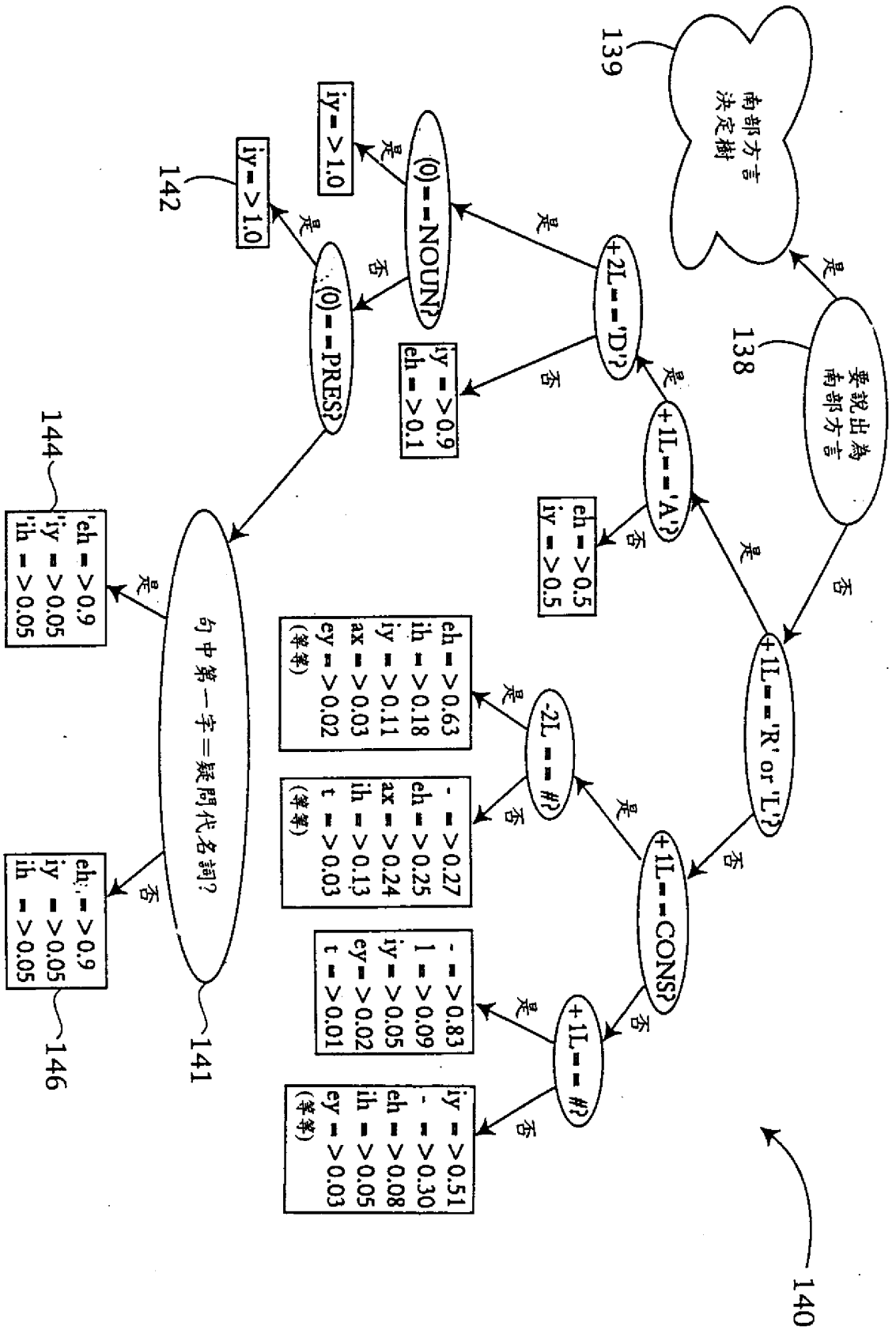


第 7 圖

( D-I-D Y-O-U K-N-O-W  
W-H-O R-E-A-D T-H-E  
A-U-T-O-B-I-O-G-R-A-P-H-Y? )



第 8 圖



第 9 圖