

國語元音聲調分析

彭玉峰*

摘要

國語語音中共有五個聲符（一聲、二聲、三聲、四聲與輕聲）；本文嘗試在元音 I 分別加上四種不同的聲調變化（不含輕聲）後，將此聲波經多尺度分析，探討聲調在不同尺度中的變化。

關鍵詞：元音、尺度、多尺度分析、聲符。

*康寧醫護暨管理專科學校 資訊管理科講師

Tones Analysis For a Mandarin Vowel

Yu-Fong Peng*

Abstract

Mandarin Chinese is a tonal language in which there are five lexical tones (Tone 1、Tone 2、Tone 3、Tone 4 and neutral tone).I made every single lexical tones except the neutral tone in an isolated syllable .The multiresolution analysis is used to analyze the vowel sound. The result indicate that tones with a variety of scales.

Keywords : multiresolution analysis 、 scale 、 tone 、 vowel.

* Lecturer, Department of Information Management,
Kang-Ning Junior College of Medical Care and Management

壹、導論

從參考文獻【1】中指出國語是一種有聲調的語言，聲調的音位是以高低的變化型態差異，在每個不同的音節上分別出四個不同的詞義。這四個聲調以注音符號來書寫，分別是「一聲：一」、「二聲：ㄨ」、「三聲：ㄨˇ」及「四二聲：ㄨˋ」；它們正表示三種特質：（1）高低程度不同，（2）時間長短不同，（3）高低變化型態的不同。

參考文獻【2】四聲辨識用的是聲調軌跡的韻取，參考文獻【3】的聲調分析則是利用四個聲調的能量來進行。本次研究基於聲調有四種變化，且語音訊號是一維訊號，站在訊號處理的方法而言，我們可以利用基底的轉換進行訊號處理，在 1988 年 Daubechies 運用多尺度分析【4】的方法製造出一族有限覆蓋（compact support）的正交基底【5】。在參考資料【6】，我們已經使用有名的 D4 係數，做為低頻濾波係數 C_k ，藉以作成低頻濾波矩陣；高頻濾波係數則經由小波方程式將低頻濾波係數 C_k ，轉換成高頻濾波係數，再作成高頻濾波矩陣。因此我們將羅馬拼音中的元音 I，分別加上四種聲調的變化後，希望經有多尺度的分析的『顯微』功能，嘗試在不同尺度中來看聲調的變化。

貳、數學分析方法

一、向量空間（Vector Space）：

任意 n 個向量 v_1, v_2, \dots, v_n 和 n 個純量 a_1, a_2, \dots, a_n ，

$$a_1 v_1 + a_2 v_2 + \dots + a_n v_n = \sum_{i=1}^n a_i v_i \quad (2.1)$$

我們將式（2.1）稱為「向量 v_1, v_2, \dots, v_n 的線性組合」。

當 $X = a_1v_1 + a_2v_2 + \dots + a_nv_n$ ，我們稱爲「 X 被向量 $v_1、v_2、\dots、v_n$ 展開」。

如果 $\{v_1, v_2, v_3, \dots, v_n\}$ 是一組向量，則它所有的線性組合所成的集合爲

$$\text{span}\{v_1, v_2, \dots, v_n\} = \left\{ \sum_{i=1}^n a_i v_i \mid a_i \in \mathbf{R} \right\} \quad (2.2)$$

令 $S = \text{span}\{v_1, v_2, \dots, v_n\}$ ，我們稱「向量 $v_1、v_2、\dots、v_n$ 展開了空間 S 」

若 $\{v_1, v_2, v_3, \dots, v_n\}$ 是一組展開空間 S 的一組最精簡向量；換言之，就是這 n 個向量缺少一個就無法展開空間 S ，我們稱 $\{v_1, v_2, v_3, \dots, v_n\}$ 爲空間 S 的一組基底。

二、多尺度分析 (Multiresolution Analysis ; MRA) :

任意一個原始訊號 $f(x) \in L^2(\mathbf{R})$ 上，空間 $L^2(\mathbf{R})$ 中的多尺度分析必須滿足以下性質：

$$(一) V_j \subset V_{j+1}, \forall j \in \mathbf{Z};$$

$$(二) \bigcup_{j=-\infty}^{\infty} V_j = L^2(\mathbf{R}) \text{ 和 } \bigcap_{j=-\infty}^{\infty} V_j = \{0\}, \forall j \in \mathbf{Z};$$

$$(三) f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}, \forall j \in \mathbf{Z};$$

Gilbert Strang 在1989年的論文【7】中提出的伸張方程式(dilation equation)與小波方程式(wavelet equation)，並且就有限長度爲 N 的訊號來討論，配合參考文獻【8】的等效計算可以得到以下兩個方程式及其相對的濾波方式：

伸張方程式 (dilation equation : 相當於低頻濾波)

$$\Phi(x) = \sum_k c_k \Phi(2x - k) \quad (2.3)$$

及小波方程式 (wavelet equation : 相當於高頻濾波)

$$W(x) = \sum_k d_k \psi(2x - k) = \sum_k (-1)^k c_k \Phi(2x + k - N + 1) \quad (2.4)$$

令 $\{\phi(2^j x - k)\}$ 為 V_j 的一組基底，則

$$V_j = \left\{ f(x) \mid f(x) = \sum_{k=0}^{2^j-1} c_k \phi(2^j x - k), c_k \in R \right\} \quad (2.5)$$

從式 (2.4) 的伸張方程式及式 (2.5)；顯而易見的，因為 $\{\phi(2^j x - k)\}$ 是 V_j 的一組基底，而 $\{\phi(2^j x - k)\}$ 也可以由 V_{j+1} 的基底構成。這意味的是

$$\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset V_3 \subset \cdots \subset V_N \quad (2.6)$$

這正是 S.G.MALLAT 在 1989 年根據式(2.3)提出了多尺度分析的理論與實現的方法；他將空間的展開以巢狀來表示：**【4】**

$$V_j \subset V_{j+1}, \forall j \in Z; \quad (2.7)$$

經由式 (2.7) 可以知道下列結果：

$$\text{任意函數 } f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1} \quad (2.8)$$

又令 $\{\psi(2^j x - k)\}$ 為 W_j 的一組基底，

$$\text{則 } W_j = \left\{ g(x) \mid g(x) = \sum_{k=0}^{2^j-1} d_k \psi(2^j x - k), d_k \in R \right\} \quad (2.9)$$

如果 W_j 是在 V_{j+1} 中正交於 V_j 的函數集合，則藉由伸張方程式與小波方程式所生成的基底，基於式 (2.2) 的向量空間生成法則，可以用遞迴的方式構築出一個完備的向量空間，我們寫成下列形式：

$$V_{j+1} = V_j \oplus W_j, \forall j \in N; \text{ 其中 } V_0 \perp W_0 \perp W_1 \perp \cdots \perp W_n \quad (2.10)$$

藉著式 (2.10)，可以提供了一個尺度由粗到細，解析度由低到高，由近到遠的漸近過程。

三、數位濾波器：

STRANG AND NGUYEN定義「數位濾波器是一種線性算子」【9】；是由固定長度的濾波係數所構成，而這些濾波係數通常是按照某些條件而自行設定的一組常數；輸出訊號可以藉由濾波係數和輸入訊號做旋捲和獲得。

在參考文獻【9】提出濾波器分成以下兩種：

(一) 低通濾波器：〈STRANG AND NGUYEN，1995〉實質意義為「動態計算輸入訊號的平均數」。假設輸入訊號為 X ，低通濾波係數為 $L[k]$ ，低通濾波矩陣為 L ，低頻輸出為 Y_L ，則數學表示法如下：

$$Y_L[n] = \sum_k L[k]X[n-k] \quad \text{等效於} \quad Y_L = L X \quad (2.11)$$

(二) 高通濾波器：〈STRANG AND NGUYEN，1995〉實質意義為「動態計算輸入訊號的差分數」。假設輸入訊號為 X ，高通濾波係數為 $H[k]$ ，高通濾波矩陣為 H ，高頻輸出為 Y_H ，則數學表示法如下：

$$Y_H[n] = \sum_k H[k]X[n-k] \quad \text{等效於} \quad Y_H = H X \quad (2.12)$$

今假設 H 為高通濾波矩陣， H^T 為其轉置矩陣， L 為低通濾波矩陣， L^T 為其轉置矩陣， I 為單位矩陣， O 為零矩陣，它們滿足以下的正交條件：

$$(一) \quad H H^T = I \quad (2.13)$$

$$(二) \quad L L^T = I \quad (2.14)$$

$$(三) \quad L H^T = H L^T = O \quad (2.15)$$

經由高通、低通濾波器的定義及其正交性質，MALLAT提出了分解演算法【4】：

令 $V_j = f(x)$ ，當 $j = 2p, \dots, 0$ ，計算遞迴式

$$V_{j-1} = L V_j, \quad W_{j-1} = H W_j \quad (2.16)$$

參、研究方法

一、使用軟體

(一) 採用 ARACK32 的錄音軟體以 8bits 的樣本大小，取樣率為 22050Hz 錄置語音。語音資料為 0 至 255 的字元型態，128 代表聲音的訊號為零。

檔案格式為 wav 檔。

(二) 程式語言採用 TURBO C。

二、演算法

(一) 演算法 3 · 1 (語音訊號的分解)

輸入：原始語音字元檔

輸出：多尺度的分解訊號

1、讀入字元長度為 $N = 2^n$ 的語音資料；若資料長度不足 N ，以字元資料 128，補足長度至 N 。

2、定義高頻、低頻濾波係數

3、While ($N > 1$) do

3.1if ($N = 2$)

3.1.1.利用長度為 2 的濾波矩陣與前級訊號進行旋捲和；得到 $D[0]$ 及 $D[1]$

3.1.2. $N \leftarrow N / 2$

else

3.1.3.for $j = 0$ to $j < N / 2$ do

〈3.1.3.1〉 for $i = 0$ to $2 * j + 4$ do 與前級訊號進行旋捲和

3.1.4. $N \leftarrow N / 2$

4、得到分解係數資料； $D[1 : N]$ ；

(二) 演算法 3 · 2 (特定的尺度訊號輸出)

輸入：多尺度的分解訊號

輸出：特定尺度的訊號

1、讀入經由演算法 3 · 1 處理後，所獲得的係數資料，訊號全長定為 len 。

2、定義需要輸出的特定尺度；定義為 n 。

定義特定尺度訊號的起始位置 $front$ ；初值為 0；

定義特定尺度訊號的結束位置 $rear$ ；初值為訊號全長二分之一

3、for $i=0$ to $i<n$ do

3.1 $front \leftarrow rear$;

3.2 $len \leftarrow len/2$;

3.3 $rear \leftarrow front+len$;

4、將步驟三得到的特定尺度訊號起始位置 $front$ 和結束位置 $rear$ ，對應到輸入的係數資料，就是特定尺度 n 的分解訊號。

肆、研究結果

表 4.1 至表 4.4 所列出的結果為以羅馬拼音 I 為元音的四個聲調原始語音檔、經過演算法 3.1 的各尺度分解訊號以及演算法 3.2 的 6 個特定尺度的訊號的相關資料與圖形。

表 4.1、羅馬拼音 I 的一聲相關檔案與圖形

類 別	名 稱	波 形 圖
原始語音檔	Uf_1.wav	圖 1
多尺度訊號分解檔	Uf_1wt.wav	圖 2
W4 (最高解析度)	Uf_1s5.wav	圖 3
W3	Uf_1s4.wav	圖 4
W2	Uf_1s3.wav	圖 5
W1	Uf_1s2.wav	圖 6
W0	Uf_1s1.wav	圖 7
V0 (最低解析度)	Uf_1s0.wav	圖 8

表 4.2、羅馬拼音 I 的二聲相關檔案與圖形

類 別	名 稱	波 形 圖
原始語音檔	Uf_2.wav	圖 9
多尺度訊號分解檔	Uf_2wt.wav	圖 10
W4 (最高解析度)	Uf_2s5.wav	圖 11
W3	Uf_2s4.wav	圖 12
W2	Uf_2s3.wav	圖 13
W1	Uf_2s2.wav	圖 14
W0	Uf_2s1.wav	圖 15
V0 (最低解析度)	Uf_2s0.wav	圖 16

表 4.3、羅馬拼音 I 的三聲相關檔案與圖形

類 別	名 稱	波 形 圖
原始語音檔	Uf_3.wav	圖 17
多尺度訊號分解檔	Uf_3wt.wav	圖 18
W4 (最高解析度)	Uf_3s5.wav	圖 19
W3	Uf_3s4.wav	圖 20
W2	Uf_3s3.wav	圖 21
W1	Uf_3s2.wav	圖 22
W0	Uf_3s1.wav	圖 23
V0 (最低解析度)	Uf_3s0.wav	圖 24

表 4.4、羅馬拼音 I 的四聲相關檔案與圖形

類 別	名 稱	波 形 圖
原始語音檔	Uf_4.wav	圖 25
多尺度訊號分解檔	Uf_4wt.wav	圖 26
W4 (最高解析度)	Uf_4s5.wav	圖 27
W3	Uf_4s4.wav	圖 28
W2	Uf_4s3.wav	圖 29
W1	Uf_4s2.wav	圖 30
W0	Uf_4s1.wav	圖 31
V0 (最低解析度)	Uf_4s0.wav	圖 32

圖 1 羅馬拼音 I 一聲原始語音波形

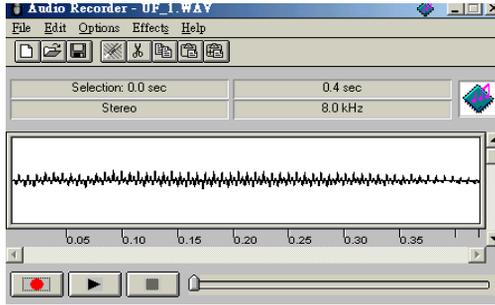


圖 2 羅馬拼音 I 一聲的多尺度分解波形

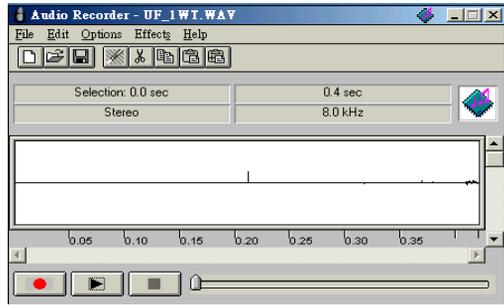


圖 3 W4 的波形

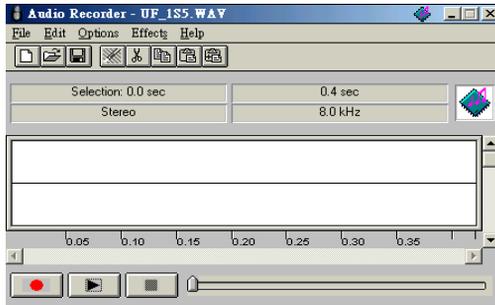


圖 4 W3 的波形

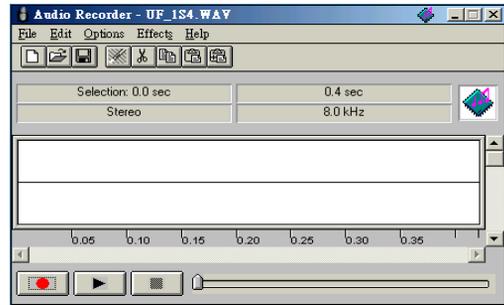


圖 5 W2 的波形圖

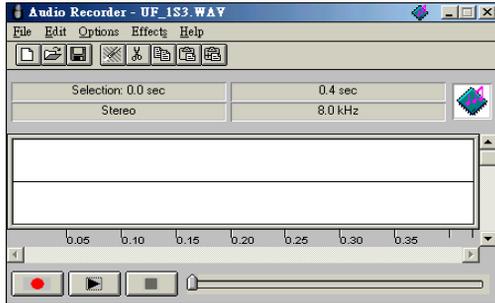


圖 6 W1 的波形圖

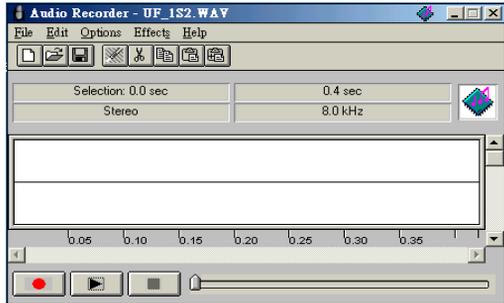


圖 7 W0 的波形圖

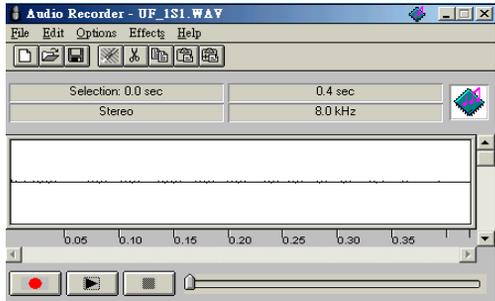


圖 8 V0 的波形圖

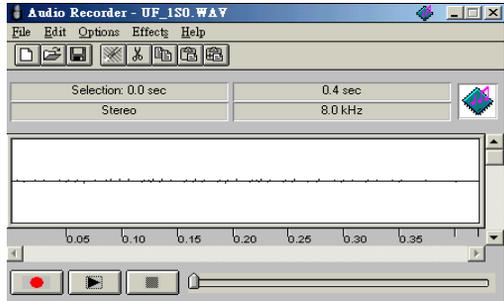


圖 9 羅馬拼音 I 二聲原始語音波形

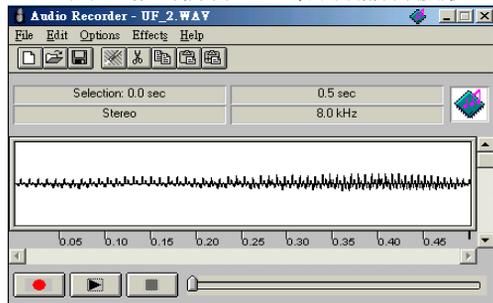


圖 10 羅馬

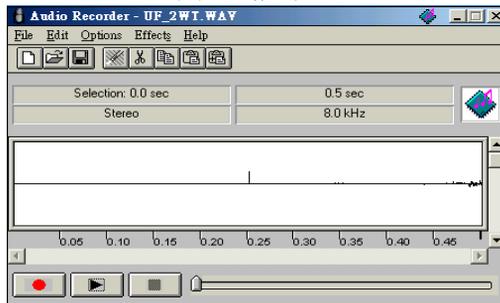


圖 11 W4 的波形

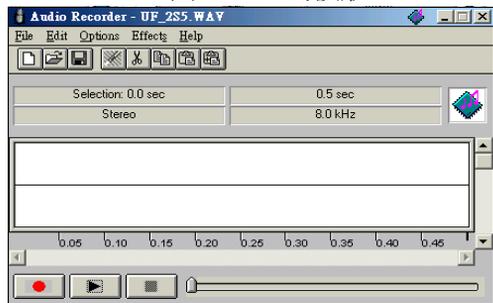


圖 12 W3 的波形

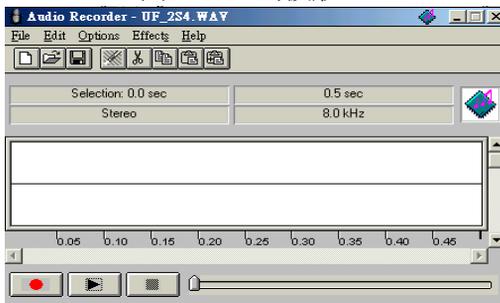


圖 13 W2 的波形圖

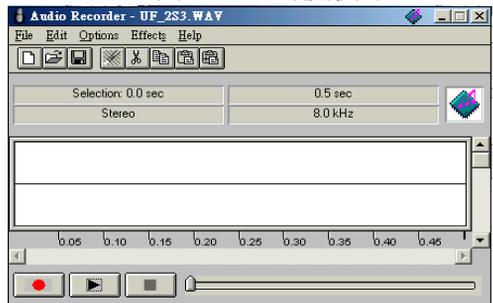


圖 14 W1 的波形圖

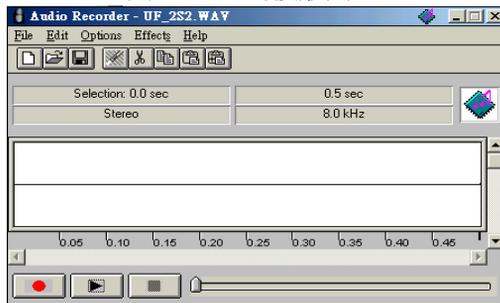


圖 15 W0 的波形圖

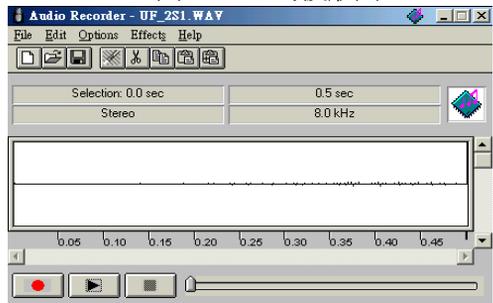


圖 16 V 0 的波形圖

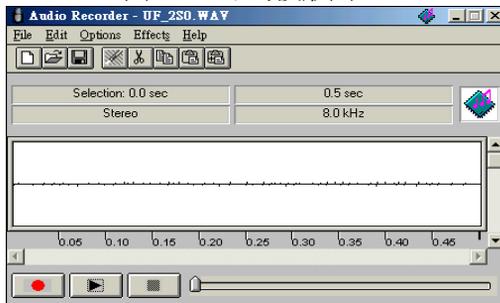


圖 17 · 羅馬拼音三聲原始語音波形

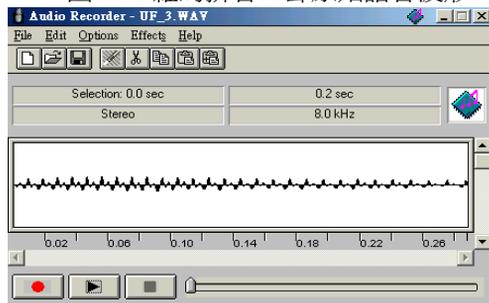


圖 18 · 羅馬拼音 I 三聲的多尺度分解波形

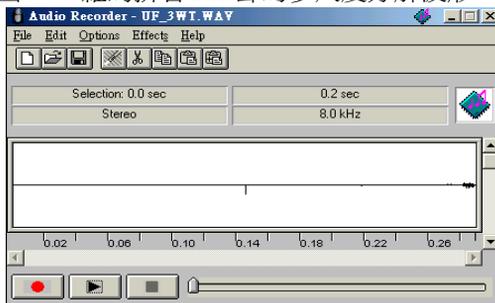


圖 19 W4 的波形



圖 20 W3 的波形

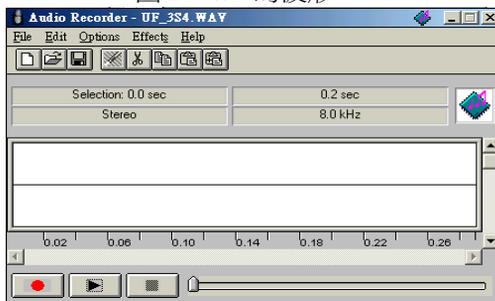


圖 21 W2 的波形圖

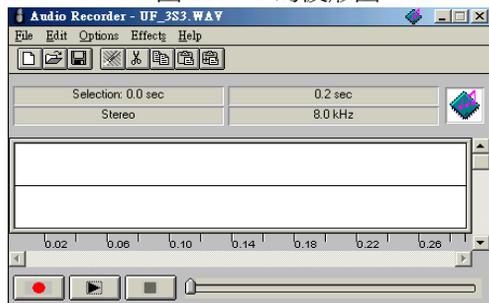


圖 22 W1 的波形圖

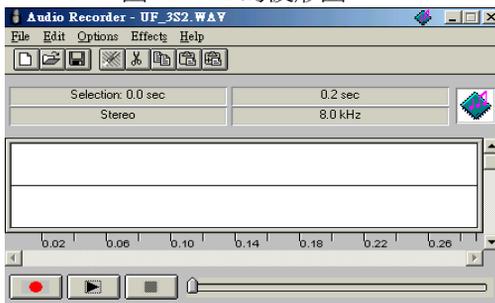


圖 23 W0 的波形圖

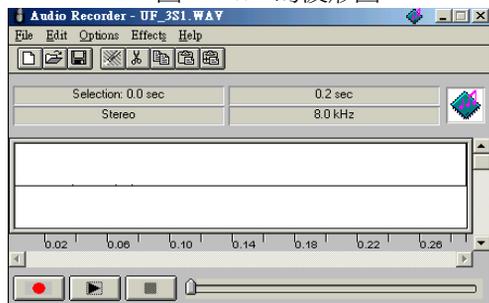


圖 24 V 0 的波形圖

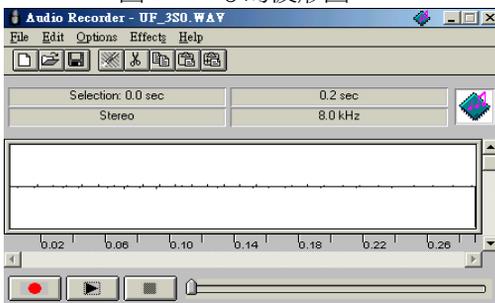


圖 25 羅馬拼音 I 四聲原始語音波形

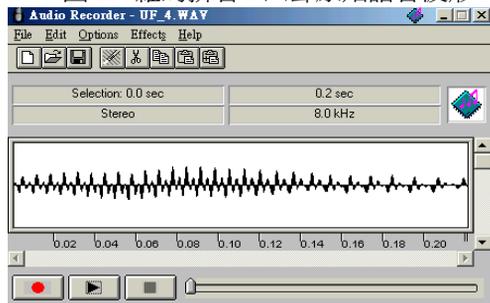


圖 26 羅馬拼音 I 四聲的多尺度分解波形

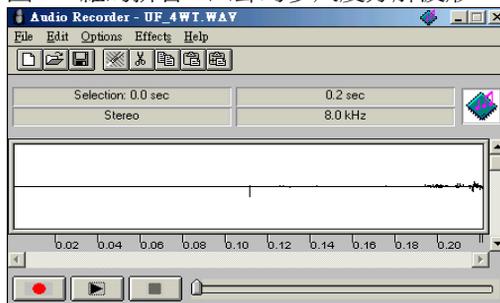


圖 27 · W4 的波形

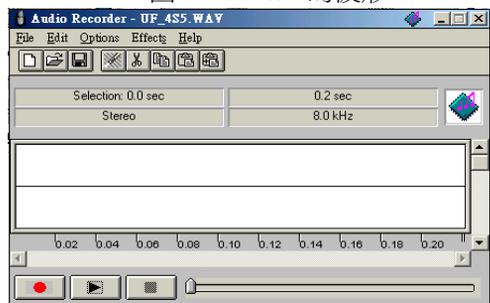


圖 28 · W3 的波形

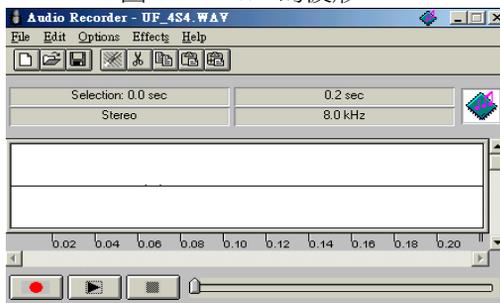


圖 29 W2 的波形圖

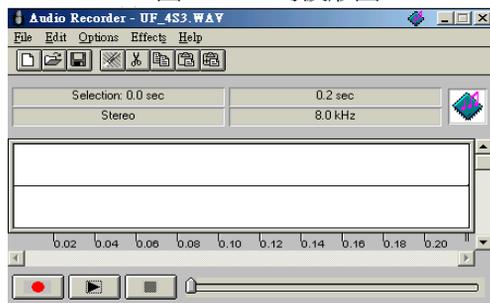


圖 30 W1 的波形圖

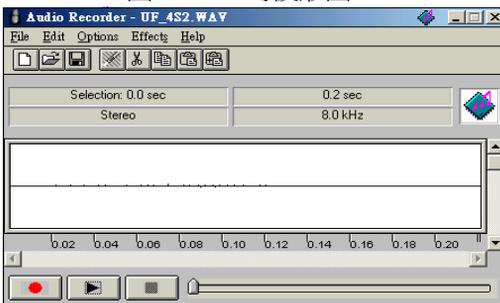


圖 31 W0 的波形圖

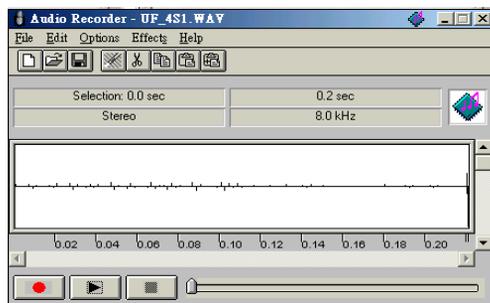
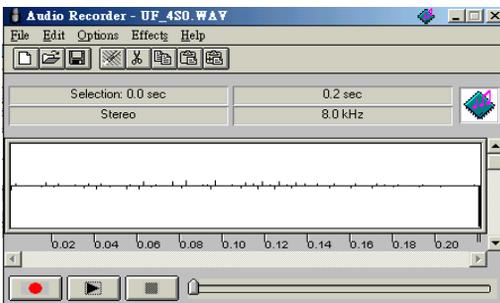


圖 32 V 0 的波形圖



伍、結論

我們從上述的研究結果中，我們可以規納出以下幾點：

- 一、我們發現較豐富的訊號出現的尺度分別是一聲和二聲出現在 W_0 和 V_0 ；三聲出現在 V_0 ；四聲出現在 W_1 、 W_0 和 V_0 。它們代表的是語音訊息在高解析度（高頻）的尺度中幾乎沒有；但在低解析度（低頻）的尺度中卻是非常豐富，我認爲這與我們人耳的構造有關，畢竟人耳接收的聲波都是低頻訊號。
- 二、從圖二和圖十，我們可以發現一聲的小波轉換訊號在後端的變化比較平緩，但二聲在後端的訊號變化就比一聲來得豐富；恰巧與一聲和二聲的特質不謀而合。
- 三、從圖十八，我們發現三聲經小波轉換後的訊號，在訊號的後端非常密集，恰與三聲的聲調起伏變化較大有關。
- 四、從圖二十六，我們發現四聲經小波轉換的訊號，在訊號的後端有持續的延伸，應該是四聲的蘊藏能量最高，所以可以在訊號衰減過程較爲持久，所以在較高的解析尺度中也會出現訊號。

這些結果真令人興奮；因爲我們可以繼續嘗試在不同的羅馬拼音元音之下，聲調是否可以一樣擁有如此豐富的結果？如果結論是肯定的，代表我們可以進行語音的切割，畢盡我們說話都是以音調作爲結尾的；這也意味著電腦未來可以從音調的變化進行語音辨認或語音合成，讓電腦的功能更進一步。

參考文獻

- 鄭傑麟，「國語聲調音位的探討」，國立屏東師範學院學報，第八期，1995，257-276—【1】。
- 陳正德等，「一種簡便的國語四聲辨認方法」，CCL RESEARCH JOURNAL，1.5，1993，36-40—【2】。
- 翁秀民等，「國語四聲的能量與字音長度之探討」，技術學刊，第十二卷，第一期，1997，125-129—【3】。
- Mallat S，「A theory for Multiresolution Signal Decomposition：The Wavelet Representation」，IEEE Trans.Pattern Anal. And Machine Intell，Vol.11，1989，674-93—【4】。
- I.Daubechies，「Ten lectures on wavelet」，SIAM，Philadelphia，1992—【5】。
- 彭玉峰，「羅馬拼音的同構數學模型」，康寧學報，第二卷，第一期，2000，1-40—【6】。
- Strang G，「Wavelet and Dilation Equation:A Brief Introduction」，SIAM Review,Vol. 31，1989，No.4，614-627—【7】。
- C.Sidney Burrus and Ramesh A. Gopinath and Haitao Guo，「Intuoduction to Wavelet and Wavelet Transforms」，Prentice Hall International Editions，1998—【8】。
- Gilbert Strang and Truong Nguyen，「Wavelet and Filter Banks」，Wellesley Cambridge，1996—【9】。