

음절신호의 음소 분리와 시간-주파수 판별 패턴의 설정

正會員 柳 光 烈*

Phoneme Separation and Establishment of Time-Frequency Discriminative Pattern on Korean Syllables

Kwang Ryol RYU* *Regular Member*

要 約 본 논문은 음절을 음소로 분리하고 각각의 특징을 추출하여 음소를 판별할 수 있는 패턴을 설정하는 실험에 관한 연구이다. 음소분리는 피치검출, 각 성문리크펄스의 폭, 포락선, 진폭의 바이어스, 발생시간 등의 파라메타를 적용한다. 최초의 피치는 성문리크의 폭, 에너지, 정규화와 성문리크의 가변바이어스 등의 변화에 따라 모음 포락선의 봉우리 부분에서 검출, 이를 기준으로 전체의 위치값을 추적한다. 모음은 포락선 패턴의 율동을 감소시키는 방법과 제2포락선만으로 모음의 판별이 가능함을 제시하며, 피치 고저에 무관한 압축파형을 추정한다. 자음은 포락선, 스펙트럼, 압축파형, 분석방법 등을 발음방법과 음소 상호 영향에서 패턴을 추출한다. 실험결과 모음음소 90%, 초성자음 80%, 종성자음 60% 판별된다.

ABSTRACT In this paper, a phoneme separation and an establishment of discriminative pattern of Korean phonemes are studied on experiment. The separation uses parameters such as pitch extraction, glottal peak pulse width of each pitch, speech duration, envelope and amplitude bias. The first pitch is extracted by deviations of glottal peak and width, energy and normalization on a bias on the top of vowel envelope. And then, it traces adjacent pitch to vowel in whole. On vowel, a method to be reduced gliding pattern and the possible of vowel distinction to be used just second formant are proposed, and shrinking pitch waveform has nothing to do with pitch length is estimated. A pattern of envelope, spectrum, shrinking waveform, and a method of analysis by mutual relation among phonemes and manners of articulation on consonant are detected. As experimental results, 90% on vowel phoneme, 80% and 60% on initial and final consonant are discriminated.

I. 서 론

인공지능형 컴퓨터에 필수적인 음성인식은 외국어에 대해 많은 연구가 진행되었고^[1,3] 우리말은 단어나 숫자음및 문장 인식의 연구가 있으며 음소분리와 음소 각각에 대한 기초적인 판별 패턴의 설정에 관한 연구도 지속적으로 진행되고

있다^[4,12].

음절 음성은 비교적 연속음에서 발생하는 조음 결합이 없고 가능한 정확하게 발음하기 때문에 고속 처리만 가능하면 실용성이 높다. 따라서 본 연구의 목적은 음성 타자기 설계를 위한 전 단계로서 음절을 음소로 분리하고 각각의 특징을 추출하여 판별할 수 있는 패턴을 설정하는데 있다. 피치검출 방법은 외국에서 다수 발표되었고^[13,14] 우리나라에서도 소수의 기법이 제시됐다^[15]. 그 중 모음의 시간파형 고찰이나 몇개의 파라메타를

*牧園大學校 電子工學科
Dept. of Electronics Eng., Mokwon Univ.
論文番號 : 91 - 125 (接受 1991. 8. 5)

적용하여 음악에 응용한 것도 있다^{16,18)}. 이들의 기법을 참고하여 모음의 포락선 최대점 부근에서 기준피치를 구하고, 이를 기준으로 전후의 피치 주기를 추적함으로써 전체의 피치구간이 설정된다. 따라서 발성시간, 진폭 비율, 바이어스 등의 파라메타를 이용하여 음소를 분리하고 음소 각각에 대해 주파수 스펙트럼 분포를 구하고 시간포락선도 추출한다. 주파수 추출은 LPC기법을 사용한다^{19,21)}. 모음에서는 포먼트 패턴의 이동과 압축(shrink) 시간파형을 추출한다. 그리고 자음에서 파열음, 마찰음, 파찰음, 비음 등의 특징, 발음방법에 따른 변화, 자 모음의 영향, 피치변화, 성대 무진동의 특징, 자음의 압축파형 등을 추정하고 중성 무성자음은 제 1포먼트를 억제하기 위해 bandstop FIR 필터를 적용^{22,23)}하여 관별 가능한 패턴을 실험으로 설정한다.

II. 음소 분리

2.1 피치주기 추적

임의의 음절신호에서 가장 큰 진폭 부분은 모음 성분으로 성분피크점들의 포락선 분포를 보면 그림 1과 같은 타원형태를 이루고있다.

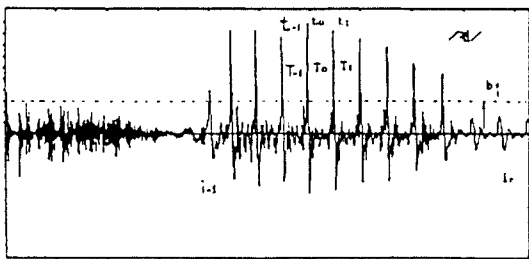


그림 1. 음절의 시간파형

이중 최고점 부분의 발성은 조음기관이 비교적 안정한 상태에서 정확하게 발음되는 순간으로 래덤성이 적기 때문에 분석 개시점으로 정한다. 그림 1의 봉우리 부분을 그림 2와 같이 확대해 고찰하면 분석 개시점에서 최고점 t_0 , 최대 진폭 A_0 와

바이어스가 0인 점이 교차하는 성분피크 펄스 정의 반파 p_0 에 대해 에너지 E_0 를 구한다. E_0 가 반이되는 값과 $0.3A_0$ 되는 높이의 에너지와 근사하다고 보면 바이어스 b_0 를 이 진폭값으로 정하여 가능한 주기내의 펄스를 적게 포함시키고 펄스 폭의 변화를 적게한다.

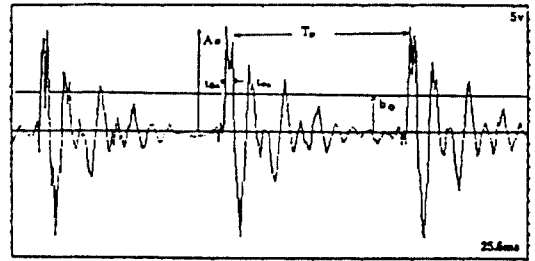


그림 2. 봉우리 부분의 파형

이때 t_0 점을 기준해 좌우로 교차하는 점 t_{up} 와 t_{dn} 에 의해 얻어지는 정의 성분 피크펄스 폭 L_0 는 식 1로 표현된다.

$$L_0 = t_{up} - t_{dn} \quad (1)$$

인접한 다른 정의 펄스 사이에 변화가 발생하므로 진폭을 에너지 값의 제곱근으로 나눠 일반화한 C_0 등과 함께 피크점 피치주기 검출을 위한 기준 파라메타가 정의된다. 따라서 p_0 와 다음에 나오는 펄스 p_m 각각의 파라메타에 대한 변화율 ϵ_i 를 계산하여 일치하는 펄스 위치 t_i 를 탐색해낸다. ϵ_i 는 실험값 $A_m=0.1$, $C_m=0.15$, $L_m=0.2$, $E_m=0.15$ 로 정한다. 이때 t_0 의 성분펄스를 Φ_0 , t_i 점의 펄스와 진폭을 Φ_i , A_i 으로 정하고 포락선 최대피크점에서의 정상피치 주기 T_0 를 식 2로 구한다.

$$T_0 = t_i - t_0 \quad (2)$$

2.2 전체 모음 구간 탐색

정상 부분의 피치주기 T_0 를 기준으로 좌우 피

치주기의 탐색으로 전체의 피치구간을 추적할 수 있다. 그림 1에서 우측방향에 대한 다음 주기 T_1 은 t_i 점에서 T_0 만큼 직접 이동한 후 $\pm T_0/2$ 부근에서 성분 피크진폭 A_0 를 찾아 Φ_2 을 얻는다. 바이아스 b_i 은 $0.3A_1$ 되는 점으로 이동하여 적응시킨다. Φ_1 와의 교차점에서 펄스 폭 L_i 를 구한다. 이와 같은 방법으로 인접된 좌우측을 탐색해 가면 식 3으로 성분펄스 Φ_i 와 피크위치 t_i 를 구하게 된다.

$$t_i = \max_{-\tau \leq t \leq +\tau} |\Phi_i(t) + T_i \pm i| \quad (3)$$

$$(-\tau = t_i - T_i - 1/2, +\tau = t_i + T_i - 1/2)$$

따라서 식 3에 의해 식 4로 피치주기 T_i 가 추출된다

$$T_i = |t_i - t_{i-1}| \quad (4)$$

$$\left(\begin{array}{l} i = -i_1, \dots, -1, 0, 1, \dots, i_1 \\ \pm i \text{에서 우측일때 } +, \text{ 좌측방향 } - \end{array} \right)$$

또한 각각의 성분펄스 폭 L_i 가 식 5로 계산된다.

$$L_i = |t_{i0} - t_{i1}| \quad (5)$$

전체의 피치주기 수는 $n = i_1 + i_2 + 1$ 이며 바이어스 b_i 에 의한 펄스 폭 L_i 의 변화로 단모음과 이중모음이 결정되는데 펄스 폭 변화율은 실험에서 약 0.15이상일때 분리된다. 그림 3의 /어/음에서 a-b까지가는 /어/음의 성분펄스 폭 변화이고 b-c는 /어/의 변화로서 분리됨을 쉽게 알 수 있다.

또한 음성 발생 시작점과 왼쪽 끝 피치의 위치 t_{i0} 와의 간격이 $t_{i0} > 2T_i$ 을 만족하면 초성이 존재함을 인식한다. 피치주기의 추적은 비음까지 탐색하므로 모음구간만의 설정은 $0.7A_1 \leq A_0$ 의 조건과 일치하지 않는 점과 음절 전체에 대해 가변 바이아스 b_i 를 $0.3A_0$ 으로 이동, $A_{i1} < A_0$ 인 점의 비교로 모음구간과 비음구간이 설정된다. 두 조

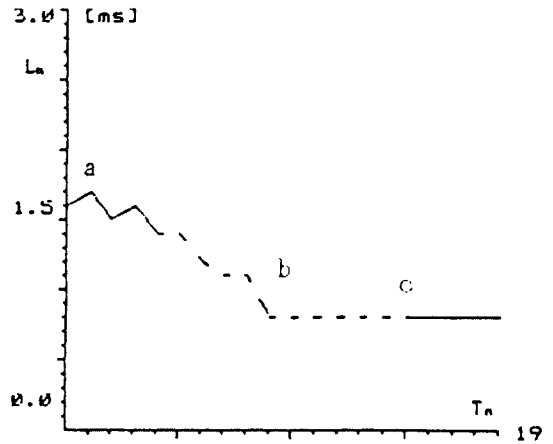


그림 3. 성분 펄스폭의 폭 변화

건을 정한 이유는 초성의 경우는 전자의 조건으로 만족하지만 중성 비음에서 모음과 유사하게 포락선이 감소하기 때문에 후자의 조건을 적용하게 된다. 무성자음은 후자의 조건에 모두 분리되거나 중성에서는 순간적으로 포락선이 감소하므로 분리 위치를 실험에 의한 결과로서 그림 6을 참고해 음절 발생시간이 80 msec 이하일 때로 정한다.

III. 시간-주파수 패턴 추출

3.1 주파수 패턴

모음 전체의 피치주기 각각에 대한 포먼트를 LPC 기법으로 구하는데 자기상관계수를 Durbin 법으로 구하는 실험예측계수를 ARMA모형(all pole model)에 적용, FFT를 취해 구한다¹²⁾. 예측계수 수는 18, 샘플링 주파수 10KHz이다. 이중모음에 대한 포먼트 변화를 보면 그림 4와 같이 앞모음과 뒷모음이 연결되는 과도부분으로서도 밀착하려는 패턴이며 유동적인 변화가 많다.

실제 /아/에서 앞/어/음이 /에/쪽으로, /아/음은 /애/쪽으로 이동하는 부의 직선함수, /와/는 정의 직선함수 형태이며 /어, 애/는 L

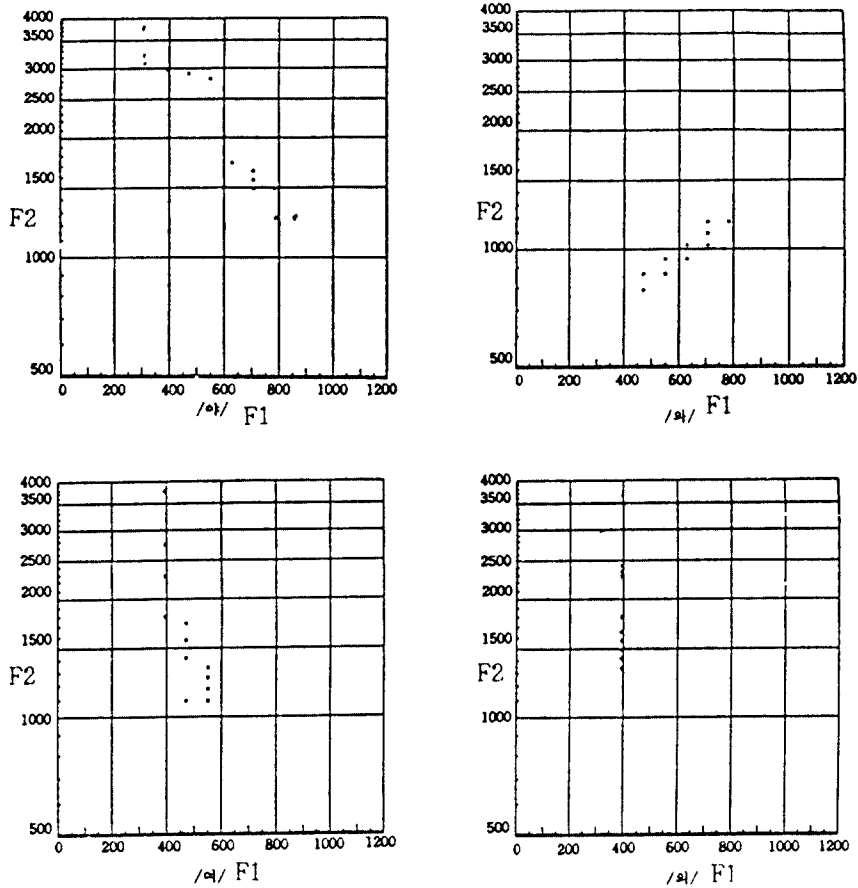


그림 4. 이중모음의 포맷트 패턴

자형, 그외는 수직 분포와 유사하다. 이에 모든 이중모음 패턴 각각을 작성해 판별해야 되지만 처리가 복잡하다. 따라서 그림 3의 성분펄스 폭 변화에서 b점 부근은 과도부분으로 뒷모음 보다 앞모음이 순간적으로 과도성이 크고 지속시간이 짧다. 또한 a점은 발음이 순간 정상발음에 도달하는 부분이며 c점은 발성이 소멸하는 위치이므로 a점에서 1.2, b점에서 2.3개 정도의 피치주기를 이동하거나 각 성분펄스 피크점의 포락선에서 앞모음은 봉우리 직전, 뒷모음은 이후 주기의 포맷트를 추출해 단모음 패턴에 적용한다.

단모음의 경우는 포락선 봉우리 즉, 안정상태의 정확한 발성 순간으로 이때의 포맷트를 구함으로써 유용성을 감소시킨다. 그림 5와 같이 8모

음 각각에 대해 고유 숫자를 지정한 표준패턴에 입력된 숫자를 overlay시켜 판별하고 새로운 측정인의 결과는 표준패턴에 보강한다. 매칭 결과가 착오영역인 ()의 경우에는 음성을 재 입력해 분석하고 그 결과에 따라 영역을 확장해 간다. 또한 피치주기와 포맷트 관계는 1.1-1.5정도의 비율이다^[21]. 따라서 표 1의 실험값에서 알 수 있듯이 4.5 msec이하의 경우는 /아/음을 기준으로 패턴을 F1축에서 1.3정도 좌측으로 이동하므로써 패턴에서 인접음성의 중첩영역을 감소시킬 수 있다. 이 음을 기준하는 이유는 모음 중에서 자연스러운 발성으로 실제 사람이 종종 발성조절에 사용하고 있기 때문이다. 따라서 실험전에 /아/음의 포맷트를 기억시켜 참고할 필요가 있다.

표 1. 피치주기와 모음의 비교

피치 주기	모음	이	이	우	에	오
남 9.6 [ms]	F1 F2 [Hz]	781 1171	312 2313	390 781	546 2031	168 703
여 4.5	F1 F2	1093 1562	468 3046	546 1015	781 2656	546 937
비		1.4 1.3	1.5 1.3	1.4 1.3	1.1 1.3	1.2 1.3

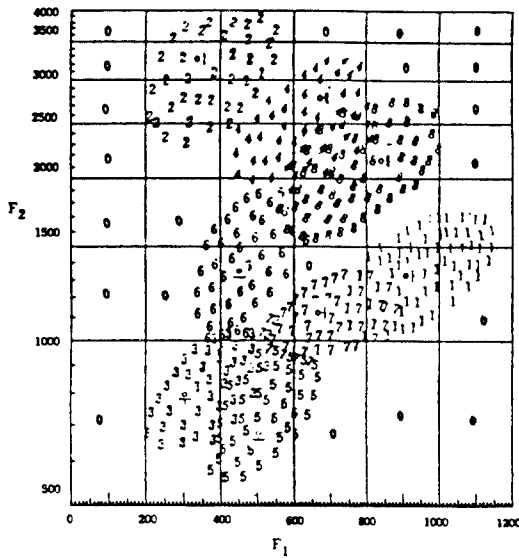


그림 5. 모음의 포먼트 배분

3.2 음절 발생시간

화자에 따라 차이가 있겠으나 그림 6에 제시한 예에서 음절발생 시간은 4명을 평균한 값으로 단모음 V를 기준할 경우 CV는 0.6, 0.8, 종성에 무성자음이면 CVC가 0.3이하, VC는 0.5, 유성자음일때 CVC는 0.6이고 VC는 0.87 정도로 종성의 무성자음이 올때 모음으로 끝날 경우 보다 반 정도의 차이로 받음된다.

초성자음의 발생시간은 그림 7에서와 같이 기성음이 가장 짧고 경성음이 연성음 보다 비교적 길다. 파열음 /ㄱ/을 기준으로 연성 경성 기성의 비율은 1:1.67:0.42, 전설음의 경우 파열음을 기준으로 파열 파찰 마찰의 비율은 1:1.75:2.5정도이다. 종성시간은 유성음을 제외하고 순

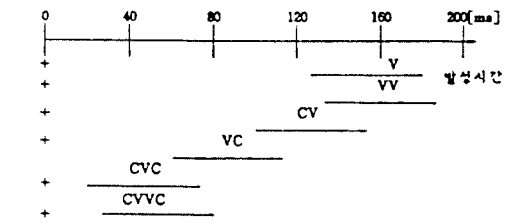


그림 6. 유성선호의 발생시간

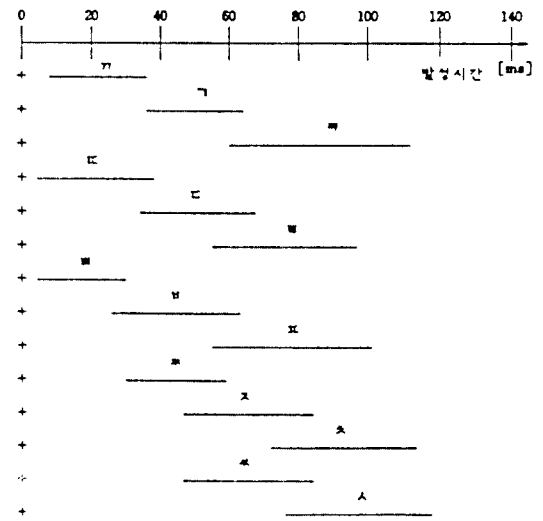


그림 7. 초성 무성자음의 발생시간

간적으로 조음기관이 닫히는 상태로 급격히 감쇄하고 그 약한 신호는 에너지가 큰 모음과 변조되기 때문에 추출하기 곤란하다. 따라서 실험값에서 20 msec 이하로 본다.

3.3 압축파형

음성의 고서는 피치주기의 장 단으로 발음방법

에 따라 또는 화자의 상태에 따라 변화한다. 그러나 어떤 상태라도 인식가능한 최소한의 파형은 존재해야 하므로 그림 8과 같이 임의의 한 주기를 구성하는 내부파형의 분석에서 3번째 내부주기까지가 판별가능한 포먼트 성분이 포함된다고 본다. 이는 중속화자의 경우 음의 고저를 점차 증가해 볼때 3까지의 파형은 변하지 않고 나머지 부분만 장 단이 이루어진다. 따라서 피치주기의 장단은 모음의 판별에 무관함을 알 수 있으며 압축파형의 위치는 정의 성분펄스 피크값의 0.3이상되는 펄스까지가 모음을 판별하는 최소파형으로 제시된다.

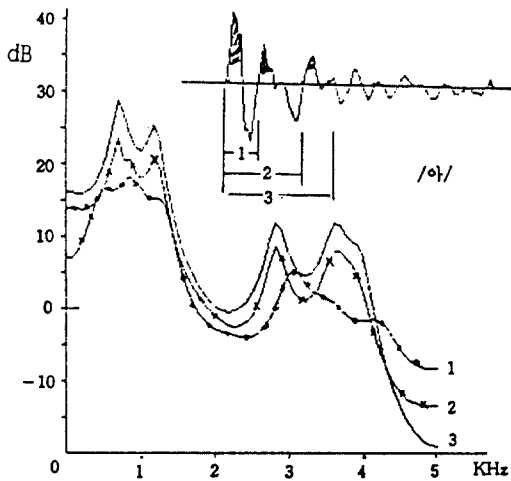
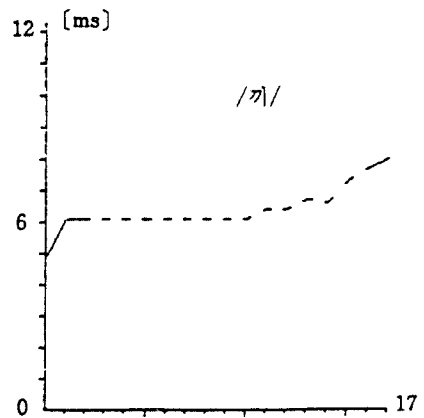
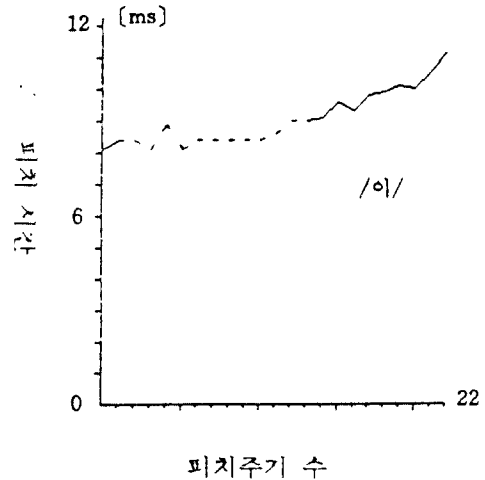


그림 8. 피치주기 내의 부분파형과 스펙트럼

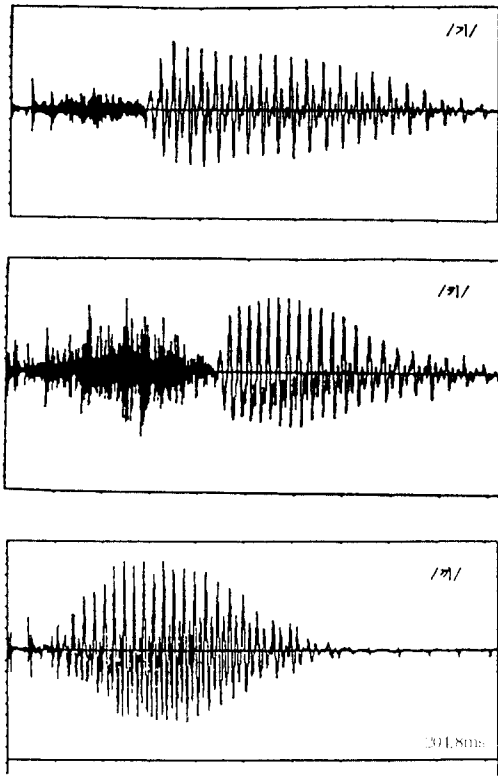
발음방법에서 피치주기는 그림 9의 a에서 초성의 연성 경성 기성에 대해 연성음 기준으로 1 : 0.7 : 0.6으로 주기가 짧아졌다가 점차 모음만의 상태로 전환된다. 예로, 보통 발성시의 주기가 8-9 msec인 특정인의 경우 /아/음에 대해 음계적으로 발음할 경우 피치주기는 옥타브 내에서 9.6, 8.7, 8.5, 7.6, 6.5, 5.9, 5.3, 5 msec 정도로 변화하는데 피치주기를 짧게했다가 어느 점에서 순간 길게 발음하면 /ㅎ/음을 포함하는 무성자음의 도래로 볼 수 있다. 여기서 단모음과 경성의

차이는 약 0.7정도이다. 이때 기성음은 자유부분이 다른 음 보다 에너지가 크게 집중되며 경성음은 적고 발생시간이 최소로 압축되고 모음의 에너지가 비교적 크다(그림 9의 b).

이 사실은 파열 마찰 파찰음에도 적용되며 초성의 경성음에 나타나는 파형은 압축된 판별 패턴으로 제시된다. 그림 10의 각 특징은 /ㄱ/은 포락선이 두개의 톱날형태, /ㄴ/은 절대치 포락선이 백색잡음 형태의 직선, /ㄷ/은 타원의 종모양, /ㅁ/는 하나의 임펄스, /ㅌ/은 임펄스 진동으로 어렵된다.

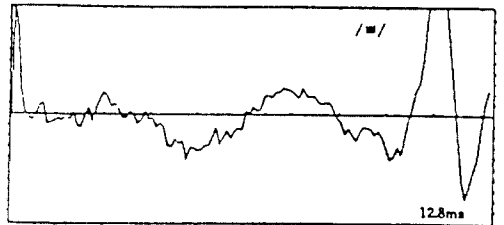
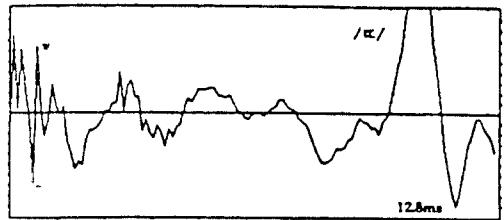
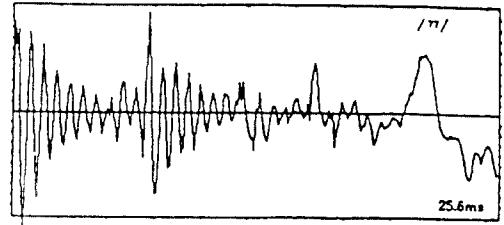


a. 피치주기의 변화



b. 언성, 기성, 장성음

그림 9. 피치추거의 변화와 발음방법의 차이



3.4 자 모음의 주파수영향

인간은 성대를 진동하지 않고 속삭임으로 말해도 인식하는 점을 분석하면 그림 11에서의 같이 제 2포먼트 성분이 가장 크게 나타난다. 성대 진동시는 에너지가 큰 제 1포먼트 주파수와 속삭임 때 나오는 음과 변조되기 때문에 일반적인 형태의 스펙트럼 분포가 된다.

이는 제 2포먼트 주파수 대역이 다른 포먼트보다 큰 환별 요인으로 작용되기 때문에 제 2포먼트 대역 만으로도 환별 가능성이 제시된다. 또한 초성자음을 발성할 때도 이미 조음기관은 초성 다음에 오는 모음의 발성 상태에서 발음되기 때문에 그림 12에서 초성자음의 주파수 대역의 분포는 성대 부 진동시에 나타난 제 2포먼트 대역만 포함된다. 이는 초성음의 주파수 추출은 환별에 크게 도움되지 못함을 알 수 있다.

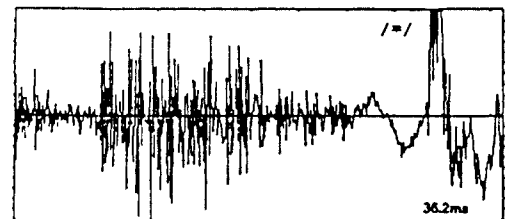
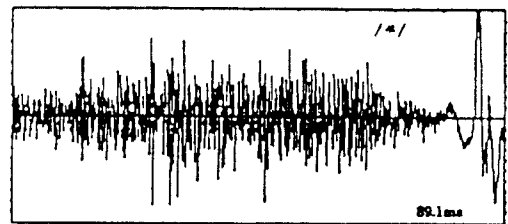


그림 10. 초성 부진자음의 입출과정

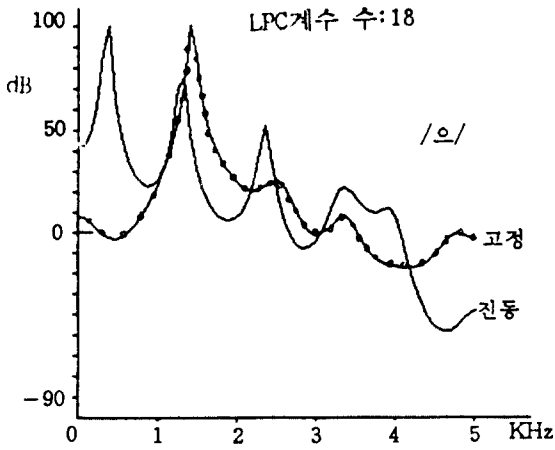


그림 11. 성대의 진동과 무 진동시의 스펙트럼

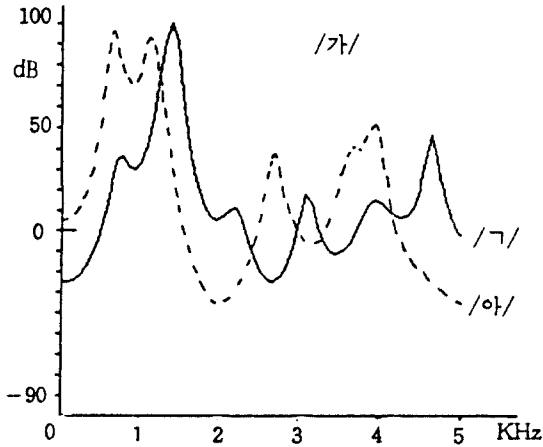


그림 12. 초성과 모음 음소의 스펙트럼

따라서 그림 13의 모음 중 일부와 같이 동일한 /ㄱ/음이라도 뒤에 오는 모음의 제 2포먼트가 /아/음 보다 /이/음이 높기 때문에 변조된 주파수만 다르고 포락선 모양은 같다. 그래서 같은 자음이라도 성대 무 진동시의 모음에 대한 성도 주파수와 변조되기 때문에 모두 다르다. 이는 순간의 시간과형 포락선 분포가 초성자음 판별에 중요한 요인이며 또한 초성 자음의 주파수 분석만으로도 중성의 모음이 판별될 수 있음이 제시 된다.

다만 자음의 주파수적인 분포는 그림 14와 같

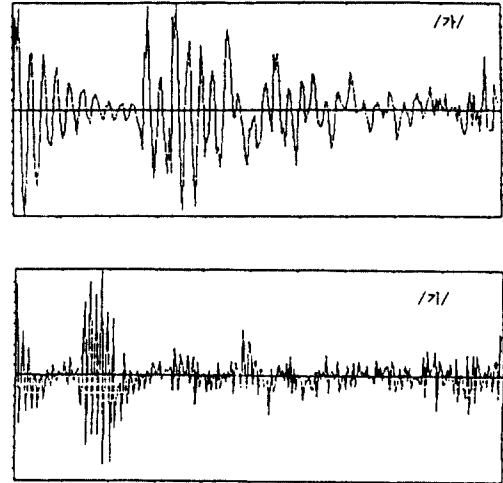


그림 13. 모음에 따른 초성자음 신호의 변화

은 패턴에서 비음, 파열, 마찰, 파찰음 영역으로 참고된다.

초성자음은 중성에 비해 발성과정이 길고 점차 열리는 상태로 구분이 용이하지만 중성자음은 비음을 제외하고 순간적으로 닫히는 동작으로 지속 시간이 짧다. 따라서 발성과정이 그림 15와 같이 초성의 역 순위로 동작하므로 서로 반대의 파형

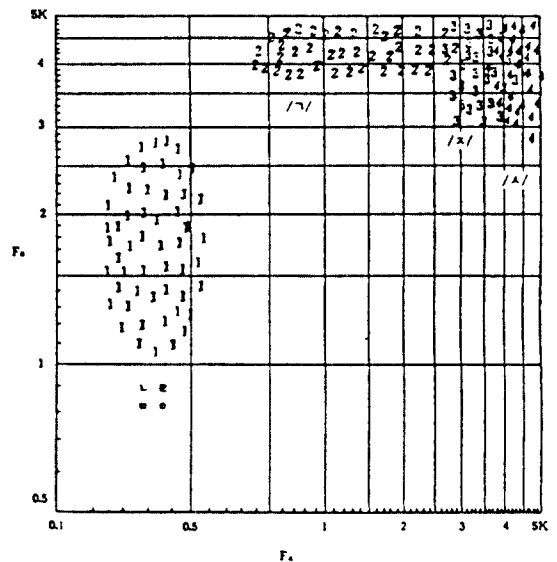


그림 14. 자음의 주파수 패턴

이 된다. 이때 자유성분은 순간적으로 에너지가 상대적으로 큰 모음과 변조하고 중성의 모음 각각에 따라 초성의 경우와 같이 다르다. 그래서 패턴 추출이 고도의 기술이 필요하지만 음절 발생 시간으로 중성 유무를 확인하고 모음 종료 과도부분의 변화에서 제 1포먼트 주파수 대역을 가능한 억제한 다음 패턴을 추출, folding하여 시간포락선을 분석한다. 사용 필터는 stopband FIR로 필터길이 50, 하한 passband는 0 400Hz와 ripple 0.002dB, 상한 passband는 모음의 제 1포먼트와 2포먼트가 만나는 골점의 주파수에서 5KHz까지로 ripple 0.002dB, rectangular window, stopband ripple 73dB의 조건으로 software에 의해 처리한다. 여기서 하한의 차단주파수를 400Hz로 정한 이유는 그림 14에서 비음을 고려했기 때문이다.

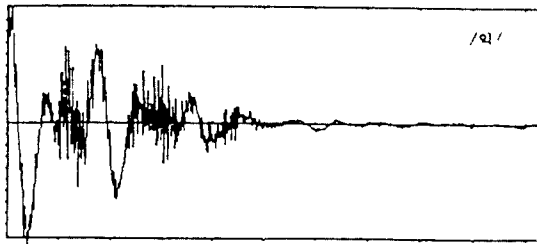


그림 15. 중성 음성자음의 변조과정

IV. 실험 및 결과

한 음절만을 분석하기 위해 대이나 회음 장치의 기능을 음성신호가 0.3 mV 이상 입력될 때 A/D 변화기의 동작이 트리거 되도록 설계한다. 이 값은 신호대 잡음비에서 백색잡음을 억제하여 인간이 인식될 수 있는 값에서 부터 입력하기 위함이다.

표 4. 초성자음 판별 결과

초성	기	니	디	기	니	스	크	크	크	크	크	크	크	크
관명수	4	2	3	2	2	4	3	3	4	4	3	4	3	4
오차수		2	1	2	2		1	1	1		1	1		
관명률	55/68, 81%													

다. 또한 주파수 성분이 높고 신호가 작은 음소를 위해 3.5KHz이상의 고역강조필터에서 30dB 증폭한다. 변화기의 기능은 분해능이 12bit, 샘플링 주파수 10KHz, 변환속도 25 μ sec, 최대입력 전압 ± 5 V 등의이며 409.6 msec 까지 변환되지만 실제 발음길이는 204.8 msec 정도이다. 모음과 자음의 주파수 패턴은 성인 남녀 10명에 대한 결과로써 음절에서 모음의 경우 포락선의 봉우리 부분의 위치를 기준으로 구한 포먼트 패턴이고 자음은 초성자음을 기준으로 추출된 주파수 패턴으로 다중의 음절을 반복 발생한 결과이다. 또한 음절발생시간은 발생 종료부분의 진폭이 봉우리 최대 진폭의 0.3되는 시간, 중성의 유성자음의 경우는 0.3구간과 직전의 위치주기에 대한 피크폭스의 0.7되는 진폭의 차이로 구한다. 초성자음은 모음 발생에서 과도부분 직전의 영교차점까지의 시간으로 추출된 결과이다. 실제 분리와 판별의 지도는 성인 남녀 각 2명의 음절에 대해 음소분리는 위치주기가 약 4 msec 이하의 경우, 착오가 발생하지만 90%이상 성공한다. 음소의 판별에서 모음을 자동 판별, 자음은 적시로 판독하고, 이를 다시 반복한 다음 판별 패턴을 설정한다. 분

표 2. 모음 판별 결과

모음	아	이	우	에	오	우	어	에
관명수	1	1	4	3	1	4	1	3
오차수				1				1
관명률	29/32, 91%							

표 3. 이중모음 판별 결과

이중모음	아	어	요	우	의	어	에	외	위
관명수	4	4	4	4	4	4	3	2	2
오차수							1	2	2
관명률	31/36, 86%								

표 5. 종성자음 판별 결과

종성	ㄱ	ㄴ	ㄷ	ㄹ	ㅁ	ㅂ	ㅇ
판별수	3	3	2	3	2	3	2
오차수	1	1	2	1	2	1	2
판별률	18/28, 65%						

표 6. 제 2포맷으로 판별한 결과

단모음	아	이	우	에	오
판별수	4	4	2	3	4
오차수			2	1	
판별률	17/20, 85%				

리된 각 음소는 그림 16와 같은 시간 주파수영역의 특징패턴 추출과정을 통한 결과 모음 음소는 약 90% 이상(표 2, 3), 초성은 80% 이상(표 4), 종성 60% 이상(표 5) 정도 판별된다. 또한 모음에서 제 2포맷으로 주파수만으로 판별한 결과(표 6) 종속화자지는 문제되지 않으나 독립화자의 경우와 /으, 이, 우/에서 중첩부분이 발생한다. 포맷트 내역폭에서 피크점의 주파수만을 선택하지 않고 포맷트 피크의 0.7배는 구간까지 적용하면 개선된다.

V. 결 론

본 연구는 음절 인식을 위한 전 단계로서 음소로 분리하고 시간과형과 주파수 분석에서 판별 가능한 패턴을 실험적으로 설정한 결과 음소 분리의 요인은 성분펄스 폭, 진폭, 가변 바이아스, 포락선, 발생시간이며 피크진폭의 상수 0.3과 0.7은 중요하게 이용된다. 또한 주파수에서 제 2포맷트 대역은 모음과 자음 판별에 핵심적인 파라미터이며 모음에서 포맷트 패턴을 이동시킬 필요가 있고 시간과형에서는 모음의 압축과정과 자음의 발생 순간포락선 분포가 시간영역 매칭시 적용도가 높다. 또한 음소의 특징은 음성합성에도 진요하게 적용할 수 있다. 따라서 특정한 하나의 기법 보다는 시간과 주파수 영역에서 종합적인

상호관계를 적용하므로써 보다 개선된 음소의 판별이 가능하다. 앞으로 계속 판별 패턴의 경제점 설정과 종성의 무성자음에 대한 더욱 정확한 패턴 추출의 연구 및 많은 사람들을 대상으로 표준화가 필요하다. 이와 같은 부분적인 보강과 고속의 매칭 알고리즘을 추가하면 충분히 음성 타자 시스템에 적용할 수 있다고 본다.

참 고 문 헌

1. Ronald W.Schafer and John D.Markel, Speech analysis, IEEE press, 1979.
2. N. Rex Dixon and Thomas B.Martin, Automatic speech and speaker recognition, IEEE press, 1979.
3. 新美康永, 音聲認識, 共立出版, 1967.
4. K.T.A., 한국어 음성인식, 협성 시스템 관한 연구, 과학기술지 연구보고서, 1986.
5. 한국과학기술원, 디지털 음성처리 기술, K.T.A. 연구보고서, 1984.
6. K.T.A., 연속음성 인식기술개발에 관한 연구, K.T.A. 연구보고서, 1987.
7. 한국과학기술원, 국어 음성의 모음, 자음, 모음연쇄의 음향분석, K.T.A. 연구보고서, 1987.
8. 윤종관, 이형수, 김화란, "Recursive least square 알고리즘을 이용한 한국어 음소분류에 관한 연구", 한국음향학회지, Vol.6, No.3, pp.60-67, 1987.
9. 유정원, 양 코덱으로 주파수 특성과 상관/PSD 패턴에 의한 한국어 음소의 분류, 경희대학교, 박사학위 논문, 1988.
10. 김준현, 한희, 박규태, "동정내역 에너지를 이용한 한국어 기본단위 음정의 자동인식에 관한 연구", 한국음향학회지, Vol.19, No.3, pp.5-12, 1982.
11. 김준현, 강정호, 변용규, 유형근, 이강성, "음소단위에 의한 한국어 연속단위 자음 인식에 관한 연구", 한국음향학회지, Vol.8, No.3, pp.5-15, 1989.
12. 김일원, 황영수, 윤대희, 차일환, "음소를 이용한 한국어 음성 신호의 분석과 인식에 관한 연구", 한국음향학회지, Vol.8, No.5, pp.70-77, 1989.
13. B.Gold and J.R.Rabiner, "Parallel processing techniques for estimation pitch periods of speech in the time domain", J.Acoust. Soc. Amer. Vol.46, pp.442-448, 1969.
14. J.R.Rabiner, M.J.Cheng, A.E.Rosenberg and C.A.

- Mogonegal, "Comparative study of several pitch detection algorithm", IEEE Trans. Acoust., Speech, and Signal Process., Vol. ASSP-24, pp. 399-418, 1976.
15. 배명진, 안수길, "면적 비교법에 의한 고속 피치검출", 한국음향학회지, Vol.22, No.2, pp.101-105, 1985.
 16. Warren H.Tucker, "A pitch estimation algorithm for speech and music", IEEE Trans. Acoust., Speech, and Signal Process., Vol. ASSP-26, No.6, pp.597-604, 1978.
 17. B.S.Atal and S.L.Hanauer, "Speech analysis and synthesis by linear prediction of the speech wave", J.Acoust. Soc. Amer., Vol 50, pp.637-655, 1971.
 18. 배명진, 임대열, 안수길, "음성발생 모델로 부터의 G-peak를 이용한 음성에너지 추출에 관한 연구", 전자공학회지, Vol.24, No.3, pp.381-393, 1987.
 19. 井手花之, 牧野正三, "時間-周波數ベクトルを用いた無聲破裂音の認識", 日本音響學會誌, Vol.36, No.5, 321-329, 1983.
 20. L.R.Rabiner and R.W.Schofer, Digital processing of speech signal, Prentice-Hall, 1978.
 21. 安居院 猛, コンピタ 音聲處理, 産報出版, 1983.
 22. Abraham Peled and Bede Liu, Digital signal processing theory, design and implementation, John Wiley, 1976.
 23. E.Organ Brigham, The fast fourier transform, Prentice-Hall, 1974.
 24. 伊福部 達, 音聲タイプライタの設計, CQ出版社, 1983.

이 논문은 1989년도 문교부 지원 한국학술진흥재단의 자유공모과제 학술연구조성비에 의하여 연구되었음.



柳 光 烈(Kwang Ryol RYU) 正會員
 1954年 9月 15日生
 1975年: 光云工大 無線通信工學科 工學
 士
 1980年: 慶熙大學校 電子工學科 工學碩
 士
 1988年: 慶熙大學校 電子工學科 工學博
 士

1984年~1985年: 三星電子綜合研究所 主任研究員

1985~1988年: 牧園大學 電算情報學科 專任講師

1988年~現在: 牧園大學 電子工學科 助教授

※관심분야: Computer System, D.S.P.