

# 음소에 의한 한국어 음성의 분석과 인식

## The Analysis and Recognition of Korean Speech Signal using the Phoneme

\*김 영 일 (Ki, Y. I.)  
\*이 건 기 (Lee, G. K.)  
\*이 문 수 (Lee, M. S.)

### 요 약

한국어는 발음상의 특징과 구조에 의해서 음소별로 분리가 가능하므로, 한국어를 자음 음소, 모음 음소, 받침 음소로 나눌 수 있다.

분리된 각각의 음소들을 편자기 상관계수를 이용하여 분석하였는데, 이 때 예측 차수는 15차이다. 분석 실험에서 동일한 음소들은 그 특성이 거의 유사하였다.

한국어 단음 675개를 자음 음소, 모음 음소, 받침 음소로 각각 분리하여 인식한 결과, 각각 85.0(%) , 90.7(%) , 85.5(%) 의 인식률을 얻었고, 이 음소들을 결합시킨 단음에서는 72.1(%) 의 인식률을 얻었다. 따라서, 이와 같은 방법을 이용하여 한국어 단음을 작은 데이터 양으로 처리 시간을 단축시킬 수 있고, 더 나아가 한국어의 모든 단음, 단어, 문장 등을 인식할 수 있다.

### ABSTRACT

As Korean language can be phonemically classified according to the characteristic and structure of its pronunciation, Korean syllables can be divided into the phonemes such as consonant and vowel.

The divided phonemes are analyzed by using the method of partial autocorrelation, and the order of partial autocorrelation coefficient is 15. In analysis, it is shown that each characteristic of the same consonants, vowels, and end consonant in syllables is similar.

The experiments is carried out by dividing 675 syllables into consonants, vowels, and end consonants. The recognition rate of consonants, vowels, end-consonants, and syllables are 85.0(%), 90.7(%), 85.8(%), and 72.1(%) respectively.

In conclusion, it is shown that Korean syllables, divided by the phonemes, are analyzed and recognized with minimum data and short processing time. Furthermore, it is shown that Korean syllables, words and sentences are recognized in the same way.

## I. 서론

한국어는 각 소리마다 입의 모양과 혀의 위치에 따라 발생되는 표음적인 음소문자이다. 따라서, 모음은 입의 모양에 의해 정해지고, 자음·모음·받침으로 구성된 단음은 혀, 잇몸, 입술 등에서 먼저 자음을 발생하고, 뒤이어서 계속되는 모음의 입모양으로 변천한 후, 다시 받침의 입모양으로 변천한다. 또한 연속음의 경우에는 이러한 경우가 반복된다.

지금까지 연구된 한국어 음성에 관한 연구는 주로 숫자음과 특정 지역명(전철역)의 분석과 인식에 관한 것으로, 이들의 연구는 단어 자체의 비교에 의해서 몇 개의 특정한 단어 음성만을 인식하기 때문에 인식할 수 있는 단어수가 제한되었고, 또 인식하려는 단어수가 많아지면 대량의 기억 용량이 필요하고, 처리 시간도 비교적 많이 소요된다.

본 연구에서는 한국어의 발음상의 특징과 구조에 의해서 한국어는 쉽게 음소별로 분리가능하므로, 자음·모음·받침으로 구성된 한국어 단음을 자음의 음소와 모음의 음소 및 받침의 음소로 각각 분리하고, 분리된 각 음소들을 편자기 상관 계수를 이용하여 현재 사용되는 한국어 단음 675개를 음소법에 의해서 인식하고자 한다.

## II. 한국어 음소의 분류

성대에서 입술까지인 성도를 한 개의 음향관으로 볼 수 있는데, 이 음향관인 성도의 여러 가지 모양에 따라 각각 음향적 특성이 달라진다. 따라서, 각

소리마다 성도의 모양이 다르기 때문에 각각 다른 음향적 특성을 갖고 입술에서 외부로 방사된다.<sup>5)6)7)</sup>

그림 1은 한국어의 자음을 발음할 때의 형상을 나타낸다.

그림 1에서와 같이 한국어 자음을 발음할 때, 각각의 자음에 대해서 조음 위치와 조음 상태가 다르므로, 자음들은 각각 서로 다른 음향적 특성을 갖는다. 그러나, 동일한 자음이라도 시간성과 개인성에 의해 조음 위치와 조음 상태에 약간의 차이가 있지

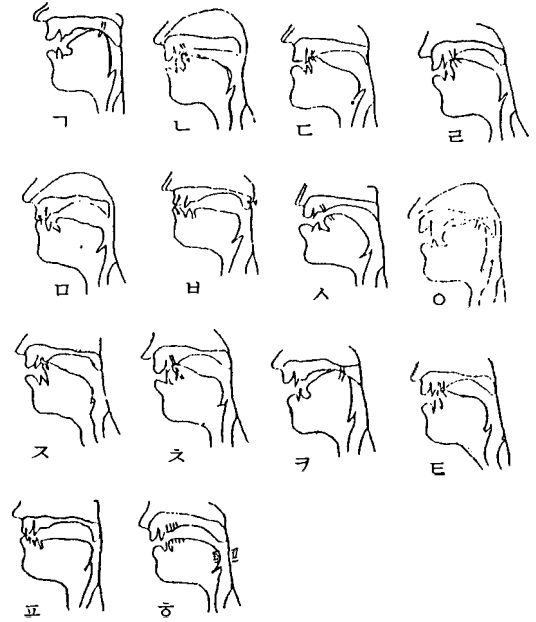


그림 1 자음 발음의 형상  
Schematic vocal tract profiles for the production of Korean consonants.

만, 동일한 자음은 조음 위치와 조음 상태가 거의 비슷하므로 그들의 음향적 특성은 유사하다.

그림 2는 한국어의 모음을 발음할 때의 형상을 나타낸다.



그림 2 모음 발음의 형상  
Schematic vocal tract profiles for the production of Korean vowels.

그림 2에서와 같이 한국어의 모음을 발음할 때, 각각의 모음에 대해서 입안의 모양이나 혀의 위치, 입술의 모양들과 같이 성도의 모양이 다르므로, 모음들은 각각 서로 다른 음향적 특성을 갖는다. 그러나, 동일한 모음이라도 시간성과 개인성에 의해 성도의 모양에 약간의 차이가 있지만, 동일한 모음은 성도의 모양이 거의 비슷하므로 그들의 음향적 특성은 유사하다.

한국어의 받침을 발음할 때에는 그림 1의 자음의 발음 형상과 거의 유사하고, 몇 개의 받침은 발음

법칙에 의한 대표되는 받침음으로 발음된다. 각각의 받침들은 조음 위치와 조음 상태가 다르기 때문에, 각각 서로 다른 음향적 특성을 갖는다. 또, 동일한 받침이라도 시간성과 개인성에 의해 조음 위치와 조음 상태가 약간 다르지만, 동일한 받침은 조음 위치와 조음 상태가 거의 비슷하므로 그들의 음향적 특성이 유사하다.

표 1과 같이 한국어 단음을 자음과 모음 및 받침들이 결합하여 이루어지고, 이러한 단음이 모여서 단어를 구성하고, 또 단어가 모여서 문장이 구성된다.

표 1 한국어 단음의 구조  
Table 1. Structure of Korean syllables

	단음의 구조	보 기
1	자음+모음	가
2	자음+모음+자음	각
3	자음+모음+모음	과
4	자음+모음+모음+자음	곽
⋮		

자음·모음·받침으로 구성된 단음을 발음할 때, 먼저 자음을 발성한 후 뒤이어서 계속되는 모음을 발성하고, 다시 받침을 발성하여 하나의 단음을 형성한다. 그러므로, 단음을 자음의 음소와 모음의 음소 및 받침의 음소로 각각 분리할 수 있다. 즉, /가/를 'ㄱ'의 자음 음소와 '아'의 모음 음소, 또 /간/을 'ㄱ'의 자음 음소와 '아'의 모음 음소 및 'ㄴ'의 받침 음소로 각각 나눌 수 있다.<sup>(8)(9)</sup>

그림 3은 한국어 단음 중 /감/, /남/, /담/, /맘/, /밤/, /삼/, /잠/, /참/, /깜/, /탐/, /팜/, /함/의 음성 파형을 나타낸다.

그림 3에서와 같이 이러한 음성 파형들은 먼저 자음 파형이 나타나고, 과도부를 거쳐 단모음 /아/의 주기적인 파형이 반복되고, 다시 과도부를 거쳐

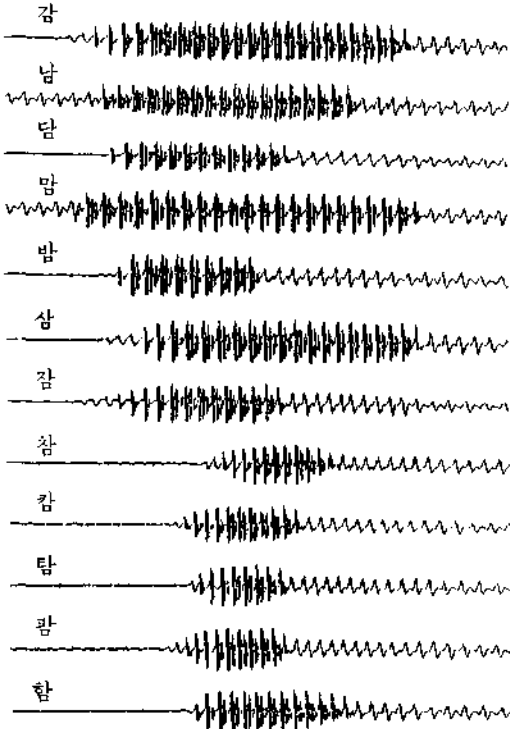


그림 3 한국어 단음의 파형  
Waveform of Korean Syllables.

받침 파형 /ㄱ/이 나타난다. 각 단음들의 파형 중에서 모음 파형은 같은 모음 음소 /아/이고, 받침 파형은 같은 받침 음소 /ㄱ/이므로 거의 유사하고, 각 음소의 파형의 길이는 발성 속도나 발성 방법에 의

해 약간 다르다.

단어나 문장을 발음할 때에는 자음과 모음 및 받침 부분의 각 음성 파형은 짧지만, 각각의 모음과 자음 및 받침 부분은 독립하여 발음할 경우와 유사하다.

그림 3에서와 같이 자음·모음·받침으로 구성된 단음의 파형에서 자음 음소와 모음 음소 및 받침의 음소로 각각 분리할 수 있다. 이 분리된 음소들을 각각 분석하여 각 음소들의 표준 패턴과 비교하여 먼저 각 음소들을 인식하고, 이 음소의 인식 결과를 결합하여 단음을 인식할 수 있다.

이상에서와 같이 한국어 단음을 인식 또는 합성을 행할 때, 음성 파형에서 자음 음소와 모음 음소 및 받침 음소를 각각 분리하여 인식 또는 합성을 행함으로써 처리할 데이터 양과 처리 시간을 줄일 수 있고, 더 나아가서 단어, 문장 등 모든 한국어를 적은 데이터 양과 빠른 처리 시간으로 인식 또는 합성할 수 있다.

### III. 실험 및 결과 고찰

#### 1. 실험 장치의 구성

그림 4는 단음을 분석하고, 인식하는 실험 장치의 전체 구성도를 나타낸다.

남자 아나운서가 보통 속도로 발음한 한국어 단음을 녹음하여 디지털 이벤트 레코더(digital event

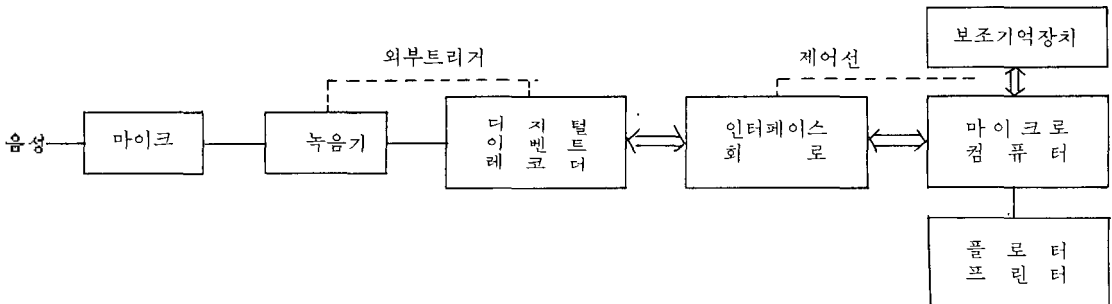


그림 4 실험 장치의 구성도  
Block diagram of experimental system.

recorder)에 입력시켜 A/D 변환을 하였다.

A/D 변환시의 샘플링 주파수는 10(KHz)이고, 8 비트로 양자화하였으며, A/D 변환의 시작 신호는 녹음기의 출력 신호의 레벨을 이용하여 외부 트리거시켰다.

디지털 이벤트 레코더에서 A/D 변환된 데이터를 인터페이스 회로를 통하여 마이크로컴퓨터로 전송시켰다. 인터페이스 회로는 PPI 칩 8255A-5를 사용하였으며, 데이터의 전송은 핸드 셰이크 방식을 사용하였다. 데이터의 전송시의 모든 제어 명령은 마이크로컴퓨터에서 프로그램으로 수행하였다.

마이크로컴퓨터에서는 인터페이스 회로의 모든 동작을 제어하고, 디지털 이벤트 레코더에서 전송된 음성 신호를 분석하여 특징 파라미터(편자기 상관 계수)를 추출하고, 이 추출된 특징 파라미터와 표준 패턴과 비교하여 음성 신호를 인식하는 과정을 수행하였다.

2. 음소의 분석

한국어 단음을 자음 음소와 모음 음소 및 받침 음소로 각각 분리하여 편자기 상관 계수로 분석한다. 이 때, 각 음소의 분리는 시각으로 행하였고, 분석한 각 음소들의 데이터 수는 각각 256개이며, rectangular window로 각 음소들의 중간 부분에서 취하였다. 예측 차수는 성도의 공진 특성인 폴만트 주파수의 추출과 밀접하게 관계 있는데, 모음의 분석에서는 예측 차수가 12차이면 충분하지만, 자음의 무성음인 경우도 고려하여 예측 차수를 15차로 한다.<sup>(1)(2)(3)(4)</sup>

그림 5의 (a)는 단음 /다/, /단/, /달/, /담/ /달/, /닷/, /당/, /닷/, /달/ 중 자음의 음소인 무성 자음 'ㄷ'을 취한 후 편자기 상관 계수로 분석한 결과를 각 차수별로 나타낸 것이고, 그림 5의 (b)는 /덕/, /넉/, /넉/, /넉/, /넉/, /넉/, /넉/, /넉/ 중 자음의 음소인 유성 자음 'ㄴ'을 편자기 상관 계수로 분석한 결과를 나타낸다. 그림 5에서 직선은 각 차

수별로 평균값을 나타낸다.

그림 5에서 보는 바와 같이, 같은 예측 차수에서 몇 개의 분석 결과만이 부분적으로 다르지만, 거의 비슷한 결과를 나타낸다.

따라서, 단음 중에서, 모음 부분과 받침 부분은 다르지만, 동일한 자음에서는 거의 같은 성질을 가짐을 알 수 있다.

그림 6은 /감/, /남/, /담/, /람/, /맘/, /밤/, /삼/, /잠/, /참/, /깜/, /탐/, /팜/, /함/ 중 모음의 음소인 '아'를 정산 상태에서 취한 후 편자기 상관

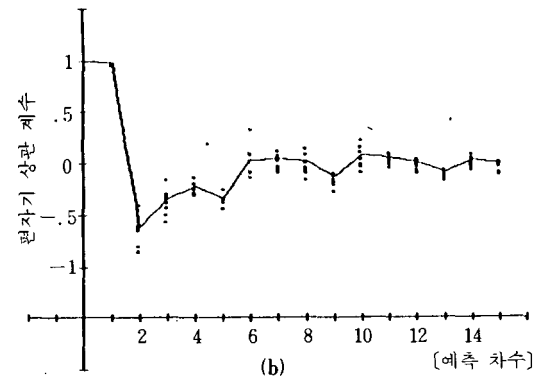
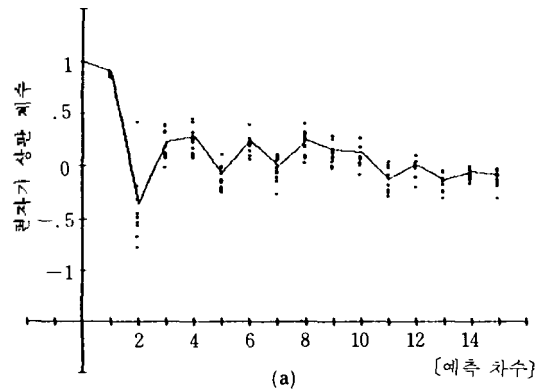


그림 5 단음 과형 중 같은 자음에서의 유사성 조사 결과  
(a) 무성 자음 (b) 유성자음  
Results of Similar Measurement of same consonant in Korean Syllables  
(a) Unvoiced (b) Voiced

계수로 분석한 결과를 각 차수별로 나타낸다.

그림 6에서 보는 바와 같이, 같은 예측 차수에 서 몇 개의 분석 결과만이 부분적으로 다르지만, 거의 비슷한 결과를 나타낸다.

따라서, 단음 중에서 처음 부분과 받침 부분은 다르지만, 모음이 같으면, 모음 부분은 같은 성질을 가짐을 알 수 있다.

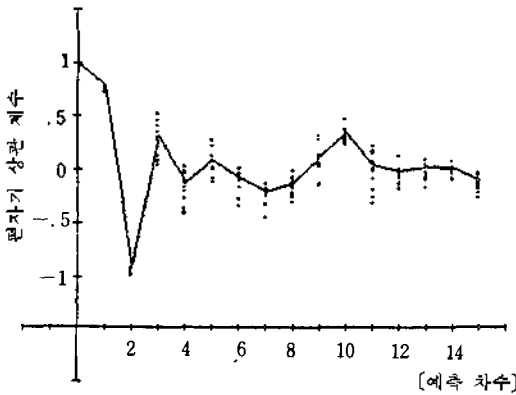


그림 6 단음 파형 중 같은 모음에서의 유사성 조사 결과  
Results of Similar Measurement of same vowel in Korean Syllables.

그림 7의 (a)는 단음 /각/, /낙/, /타/, /라/, /막/, /막/, 삭/, /작/, /착/, /탁/, /락/, /학/ 중 받침의 음소인 무성 받침 ‘ㄱ’을 취한 후 편자기 상관 계수로 분석한 결과를 각 차수별로 나타낸 것이고, 그림 7의 (b)는 단음 /검/, /넙/, /덥/, /렘/, /몹/, /빔/, /빔/, /침/, /침/, /땀/, /땀/, /힘/ 중 받침의 음소인 유성 받침 ‘ㅁ’을 취한 후 편자기 상관 계수로 분석한 결과를 각 차수별로 나타낸 것이다.

그림 7에서는 같은 예측 차수에서 몇 개의 분석 결과만이 부분적으로 다르지만, 거의 비슷한 결과를 나타낸다.

따라서, 단음 중에서 자음 부분과 모음 부분은 다르지만, 동일한 받침에서는 거의 같은 성질을 가짐을 알 수 있다.

이같은 단음을 분석한 실험 결과에서와 같이, 단음을 자음의 음소와 모음의 음소 및 받침의 음소로

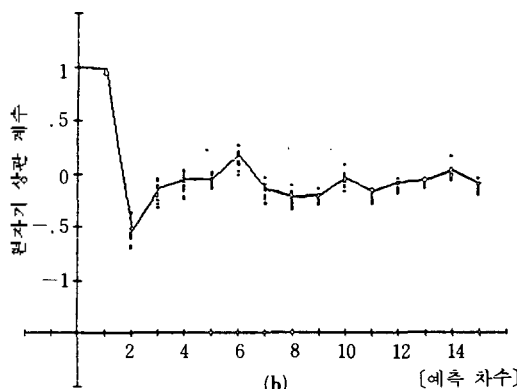
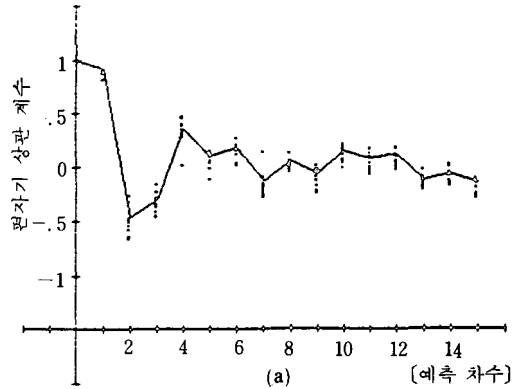


그림 7 단음 파형 중 같은 받침에서의 유사성 조사 결과  
(a) 무성받침 (b) 유성받침  
Results of similar Measurement of same end-consonant in Korean syllables  
(a) Unvoiced (b) Voiced

각각 분리하여 처리함에 있어서 동일한 자음의 음소, 동일한 모음의 음소, 동일한 받침의 음소는 그 특성이 비슷함을 알 수 있다.

그러므로, 이와 같은 성질을 이용하여 단음을 자음 음소와 모음 음소 및 받침의 음소로 각각 분리하여 인식할 수 있다.

### 3. 인식 실험

그림 8과 같이 단음에서 자음과 모음과 받침을 각각 시각으로 분리하여 네이터의 진폭을 정규화하고, 특정 파라미터인 편자기 상관 계수를 구하여 표준 패턴과 비교하여 가장 유사한 것을 판정하여 단음을 화자 중속으로 인식한다.<sup>(1)(2)(3)(4)</sup>

유사도는 다음과 같이 유클리드 거리를 이용한다.

$$\|IM - IU\| \approx \sum_{i=1}^N |m_i - u_i|$$

(1)

여기서 IM은 표준 패턴, IU는 미지 패턴을 나타내고, 표준 패턴과 미지 패턴 사이의 최소 거리를 가지는 패턴을 인식된 패턴으로 간주한다.

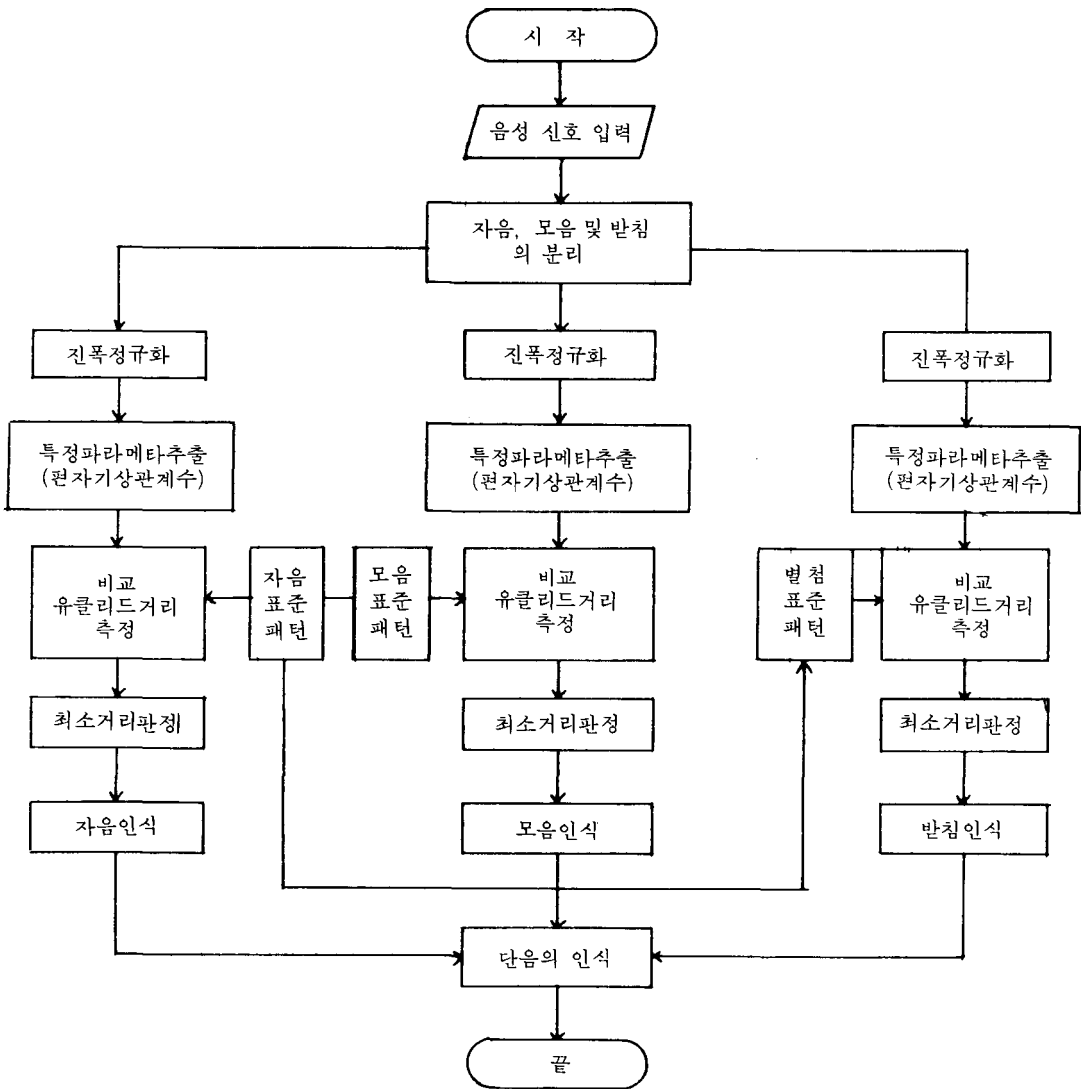


그림 8 단음의 인식 과정.  
The Flowchart recognizing Korean syllables.

표 2 단음 중에서 자음의 인식 결과  
Recognition Results of consonant in Korean Syllables

	ㄱ	ㄴ	ㄷ	ㄹ	ㅁ	ㅂ	ㅅ	ㅇ	ㅈ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ	
ㄱ			1			1				1	3	1			
ㄴ	1			4	2	1									
ㄷ		1				1	2							2	
ㄹ		3	1					1							
ㅁ		1		1		1	1								
ㅂ	2		1	2							1	2			
ㅅ			1						3	1		2			
ㅇ				4			1								
ㅈ						1	2	1		4					
ㅊ						1	5		4						
ㅋ	3			1		1	1					1	1	2	
ㅌ							4			2					1
ㅍ							2		1		1	1			2
ㅎ	3					2						1	3		
오차수	9	5	4	12	2	9	18	2	8	8	5	8	6	5	101
총 수	66	58	57	43	49	49	54	54	54	47	27	35	41	41	675
인식수	57	53	53	31	47	40	36	52	46	39	22	27	35	36	574
인 식 륜 = $\frac{574}{675} \times 100 = 85.0(\%)$															

인식 과정은 먼저 모음을 인식하고, 자음은 계속 되는 모음에 약간 영향을 받고, 또 받침은 앞의 모음에 약간 영향을 받으므로, 모음에 따라 자음과 받침의 표준 패턴을 선정하여 자음과 받침을 인식한다. 인식된 자음과 모음과 받침을 결합하여 단음(음)을 인식한다. 인식에서의 데이터 수는 256개이고, 편자기 상관 계수의 예측 차수는 15차이다.

차음 14개(ㄱ~ㅎ), 모음 6개(아, 어, 오, 우, 으, 이), 받침 14개(ㄱ~ㅎ)를 서로 조합에 의해 현재

사용되는 자음·모음과 자음·모음·받침으로 구성된 단음 675개를 편자기 상관 계수로 인식된 자음 음소와 모음 음소 및 받침 음소의 인식 결과는 각각 표 2, 표 3, 표 4와 같고, 이들 음소의 인식 결과를 결합한 단음의 인식 결과는 표 5와 같다.

받침의 인식에서는 말을 범칙에 의해서 ㄱ은 ㄱ으로, ㅅ, ㅈ, ㅋ, ㅌ, ㅎ은 ㄷ으로, ㅍ은 ㅂ으로 각각 대표 받침으로 인식을 하였다.

실제로 단음의 인식에서는 자음, 모음, 받침의 오



표 3 단음 중에서 모음의 인식 결과  
Recognition Results of Vowel in Korea Syllables

	아	어	오	우	으	이	
아		3					
어	8						
오	2	8		10		1	
우			14		11		
으						4	
이					3		
오차수	10	11	14	10	14	5	64
총 수	135	121	109	102	98	110	675
인식수	125	110	95	92	85	105	612
$\text{인식률} = \frac{612}{675} \times 100 = 90.7(\%)$							

인식의 함만큼 오차수가 증가하므로 인식률이 떨어지지만, 인식과정에서의 음소의 오인식이 중복되었기 때문에 오인식률이 높게 나타났다.

표 5 단음 인식 결과  
Table 5 Recognition Results in Korea Syllables

	자 음	모 음	받 침	단 음
인식률(%)	85.0	90.7	85.5	72.1

### IV. 결 론

한국어 단음을 자음 음소와 모음 음소 및 받침 음소로 각각 분리하여 인식하는 실험을 행하였다.

먼저 각 음소들의 분석 실험을 통하여 동일한 음소들은 그 특성이 거희 유사하므로, 한국어는 음소

표 4 단음 중에서 받침의 인식결과  
Recognition results of end-consonant in Korean syllables

	ㄱ	ㄴ	ㄷ	ㄹ	ㅁ	ㅂ	ㅅ	ㅇ	ㅈ	ㅊ	ㅋ	ㅌ	ㅍ	ㅎ	
ㄱ						7	3						1	4	
ㄴ				3	1	14									
ㄷ	14				1										
ㄹ		1													
ㅁ	1	4		1				6							
ㅂ	10	1					4		1		2	1		1	
ㅇ		1			1										
오차수	25	7	0	4	3	21	7	6	1	0	2	2	7	1	86
총 수	74	72	30	77	80	71	55	69	18	8	3	11	13	11	592
인식수	49	65	30	73	77	50	48	63	17	8	1	9	6	10	506
$\text{인식률} = \frac{506}{592} \times 100 = 85.5(\%)$															

별로 분리가 가능함을 알 수 있었다.

자음·모음과 자음·모음·받침으로 구성된 한국어 단음 675개를 자음 음소, 모음 음소, 받침 음소로 각각 분리하여 편자기 상관 계수로 인식한 결과, 각각 85.0(%), 90.7(%), 85.5(%)의 인식률을 얻었고 이 음소들을 결합시킨 단음에서는 72.1(%)의 인식률을 얻었다.

지금까지 연구된 한국어의 인식에 대한 연구는 단어 자체의 비교에 의한 숫자음과 지역명만을 인식하였지만, 본 연구에서는, 한국어 단음을 음소인 자음 음소, 모음음소, 받침 음소로 분리하여 작은 데이터 양으로 처리 시간을 단축시켜 한국어의 모든 단음, 단어, 문장 등을 인식할 수 있음을 알았다.

## 참 고 문 헌

1. R.W. Schafer & J.D. Markel, *Speech Analysis*, IEEE Press, 1979.
2. N.R. Dixon & T.B. Martin, *Automatic Speech & Speaker Recognition*, IEEE Press, 1979.
3. L.R. Rabiner & R.W. Schafer, *Digital Processing of Speech Signal*, Prentice-Hall Inc., 1978.
4. J.D. Markel & A.H. Gray, Jr., *Linear Prediction of Speech*, Springer-Verlag, 1976.
5. Gunnar Fant, *Speech Sound and Features*, MIT
6. James L. Flanagan, *Speech Analysis Synthesis and Perception*, Springer-Verlag, 1972.
7. 차일환, *음향 공학 개론*, 한신문화사, 1976.
8. 김영일, "한국어 단음의 분석과 인식에 관한 연구," 박사학위논문, 연세대학교 대학원, 1984. 12.
9. 김영일, "음소를 이용한 한국어의 인식," *한국음향학회지*, Vol. 3, No. 2, p35~p45, 1984.