

韓國語 摩擦音 및 破擦音의 分析

이 정 석 재 , 정 현 일
(영남대학교 전자공학과)

The Analysis of Korean Fricatives and Affricates

Seok-Jae Jeong and Hyun-Yeol Chung
(Dept. of Electronic Engineering, Yeung-Nam Univ.)

* Abstract

This paper describes the analysis of fricatives and affricates in Korean mono-syllables.

From the duration, averaged pattern and Fisher ratio of the consonants, we found some effective informations (optimum length) and position of features etc.) for recognition.

I. 서 론

Man-machine interface의 문제가 현실적으로 부각되고, 이를 수행하기 위한 수단으로 음성인식과 합성에 대한 연구가 활발하게 진행되고 있다.

최근 선진국에서는 국가적 차원의 음성 정보 처리 연구를 활발히 진행하고 있으며, 우리나라에서도 數字음성, 單語음성, 母音 및 子音音素, 대용량 單語 認識^[1]에 있어서 활발한 연구가 진행되고 있다.

한국어는 음소로서 단모음 10개, 자음 19개로 구성되며, 이들의 조합에 의해 실제 약 1000개의 단음절이 사용되고 있다.^[2] 또한, 한국어의 어휘는 명사의 약 90%, 용언어간의 약 86%가 1음절 및 2음절로 구성된다.^[3]

따라서, 인식의 기본 단위로서는 음소 혹은 그 이하가 유리할 것으로 생각되며, 이 경우 각 음소의 성질을 정확하게 이해하고 있지 않으면 인식시스템 설계시 발생하는 誤認識의 규명이 곤란하게 된다.

자음중 硬音 /ㄱ, ㅋ, ㆁ, ㆁ, ㄷ, ㅌ, ㅌ, ㅌ, ㅌ/은 외국어에서는 軟音 /ㄱ, ㄷ, ㅌ, ㅌ, ㅌ, ㅌ, ㅌ, ㅌ/과 구별하고 있지 않으나, 우리말에서는 그 구별이 분명하며 일상 회화에서 빈번히 사용되고 있는 한국어 특유의 음소이다.

음성인식용 전제로 한 자음의 분석과 인식에 대한 예는 破擦音 9개에 대한 鄭^[4]의 연구에서 찾아볼 수 있으나, 破擦音, 摩擦音, 鼻音, 流音에 대한 연구는 아직 불충분한 상태에 있다.

따라서, 본 논문은 음소를 인식의 기본단위로 하는 소규모 음성인식 시스템을 구현하기 위한 기초연구로서 아직까지 그 성질이 명확히 밝혀지지 않고 있는 摩擦音(/ㅅ, ㅆ, ㅎ/)과 破擦音(/ㅈ, ㅊ, ㅌ/)의 特徵 分析을 행하고 각 음소군 내에서의 식별에 유효한 Parameter들을 추출하고자 한다.

II. 韓國語 子音의 分類

한국어 자음을 조음위치와 조음양식에 의해 분류하면 調音位置別로는 響音, 軟音, 硬音, 氣音으로, 調音樣式別로는 破擦音, 摩擦音, 破擦音, 鼻音, 流音으로 분류되어지며,^[2] 그림 1과 같은 조직을 가진다.

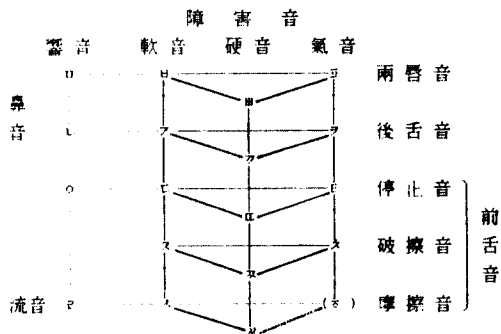


그림 1. 韓國語 子音의 組織

發聲에 있어서의 摩擦音 및 破擦音의 특징은 마찰음은 조음 기관의 어느 부분이 좁혀져서 공기의 통로가 좁아질 때 그 사이를 지나는 공기가 일종의 마찰적 운동을 일으키며 발생하는 소리이고, 파찰음은 혀바닥과 硬口蓋 사이가 폐쇄되었다가 서서히 마찰작용을 일으키며 나오는 소리이다.

III. 音聲 資料

단음절 자료로서는 한국어 逆順 辭典中の 단어를 出現 頻度順으로 나열하여 累積頻度 90% 이내에 들어가는 501개의 單音節과 음소의 수가 적은 /ㄱ, ㅋ/에 대해서는 누적빈도 99.9%까지에 들어가는 單音節로부터 뽑은 48개를 추가한 549개로 구성한다.

마찰음 및 파찰음의 단음절 음성 자료는 이중에서 선택된 160개를 방음실에서 한국인 성인 남성 3인이 3회 발생한 마찰음 720개, 파찰음 720개로 한다.

각 音素別 資料의 數는 표 1과 같다.

IV - 1. 持續時間의 分布

표 2에 각 음소에 대한 지속 시간의 평균과 標準偏差를, 그림 4에 각 音素群 別 持續 시간 分布를 나타내었다.

音素	資料數	音素	資料數
ㄷ	342	ㅅ	342
ㅈ	162	ㅊ	162
ㅊ	216	ㅎ	216

표 1. 각 音素 別 資料의 數

음소	자료수	평균 편차		음소	자료수	평균 편차	
		C+V	C+V+C			C+V	C+V+C
ㄷ	342	180.36	170.33	ㅅ	342	82.17	83.18
ㅈ	162	174.42	157.53	ㅊ	162	62.15	58.12
ㅊ	216	122.34	109.30	ㅎ	216	146.24	127.21

표 2. 지속 시간의 평균과 표준편차 (단위:ms)

IV. 分析 方法

音聲資料는 그림 2에 보인바와 같이 4.5kHz LPF를 통과한 후, 10kHz 12bit A/D 변환하여 29CH BPF (Q=6, 1/6 octave 250 Hz - 6300 Hz)를 통과시켜 분석된다. (frame 길이:10mS, 분석 window 길이: 20mS). 각 음소에는 始端, 中心, 終端 frame 이라는 시간적 label 을 부여한다.

이 Label을 참고하여 각 frame別 21차원 LPC Cepstrum 계수를 추출하여 音素別 데이터 베이스를 구성한다.

그림 3에 Labeling의 一例를 보인다.

분석에 있어서는 각 음소의 지속시간 및 평균패턴, 音素 상호간의 分散比를 통하여 각 파라미터의 인식에 대한 貢獻度, 인식에 필요한 frame의 길이등을 조사한다.

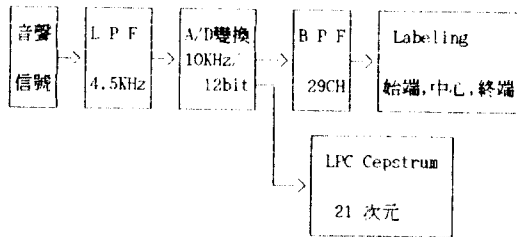


그림 2. 分析의 흐름圖

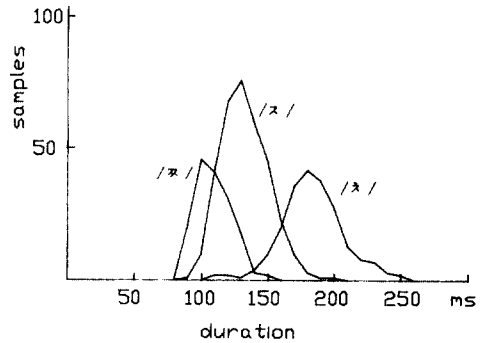
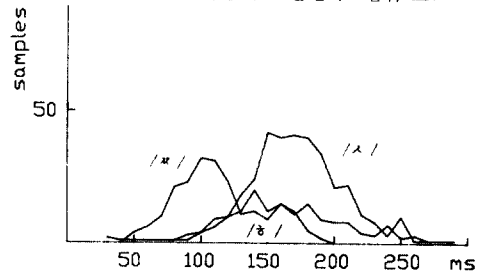


그림 4. 各 音素群 別 持續 時間의 分布

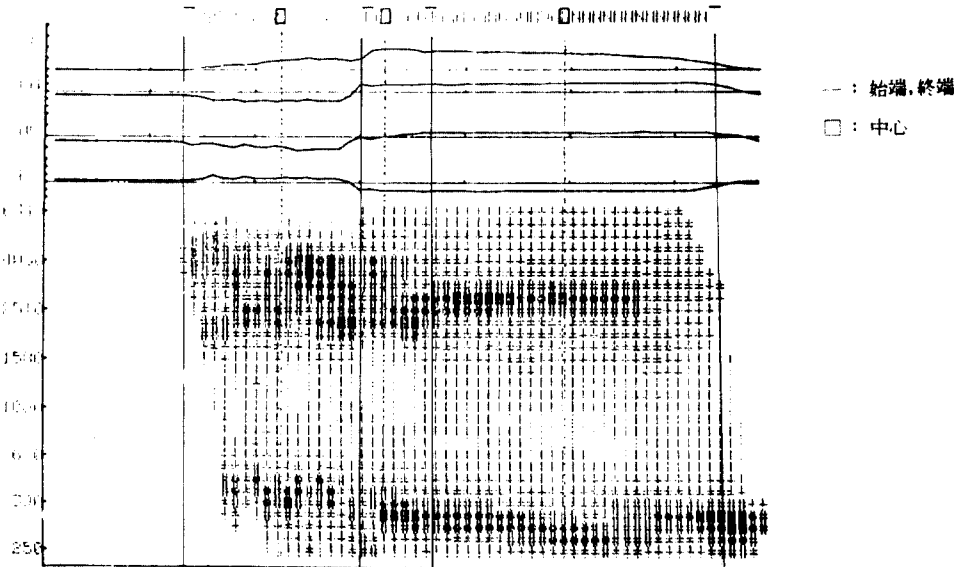


그림 3. Labeling의 一例 (單音節 "신"인 경우)

그림 4로부터 摩擦音의 경우 硬音 /ㄱ/은 軟音 /ㄷ/과 다소 重疊되지만 /ㅅ/보다 持續時間이 짧아 분리 가능하다. 氣音 /ㅎ/은 硬音, 軟音과 대부분 重疊이 되므로 지속시간만으로는 분리하기 힘들 것으로 생각된다.

또, 破擦音의 경우 氣音 /ㅈ/은 硬音 /ㄱ/과 거의 分離되어 있고 軟音 /ㄷ/과도 다소 중첩이 되지만 분리가 가능함을 알 수 있다.

따라서 이 持續時間 情報는 摩擦音에서는 軟音과 硬音이 분리예, 破擦音에서는 硬音, 軟音으로부터 氣音의 분리에 유효하다고 생각된다.

IV - 2 . Cepstrum 계수의 평균 패턴 분포

각 frame別 LPC Cepstrum 계수의 평균패턴을 조사함으로써 각 음소의 특징을 알 수 있다.

그림 5에 제 1 frame(음성 始端으로부터 첫 10 ms 구간)과 제 2 frame에서의 평균패턴 (後續 母音이 /ㅈ/인 경우)의 예를 보인다.

이로부터, 摩擦音 群과 破擦音 群의 분리가 가능하며, 摩擦音 내부에서는 /ㅅ, ㅎ/와 /ㄱ/를 분리할 수 있음을 알 수 있다.

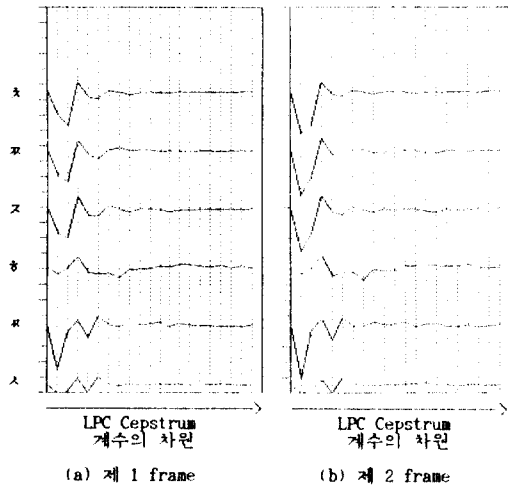


그림 5. Cepstrum 계수의 평균패턴 분포의 예 (후속모음이 /ㅈ/인 경우)

IV - 3 . power의 평균 패턴 분포

그림 6에 power의 평균패턴 (후속모음이 /ㅈ/인 경우)을 보인다.

摩擦音이 破擦音보다 始端의 power가 작게 나타남을 알 수 있다. 따라서, power成分의 평균패턴을 조사함으로써 摩擦音 群과 破擦音 群의 분리가 가능하리라 생각된다.

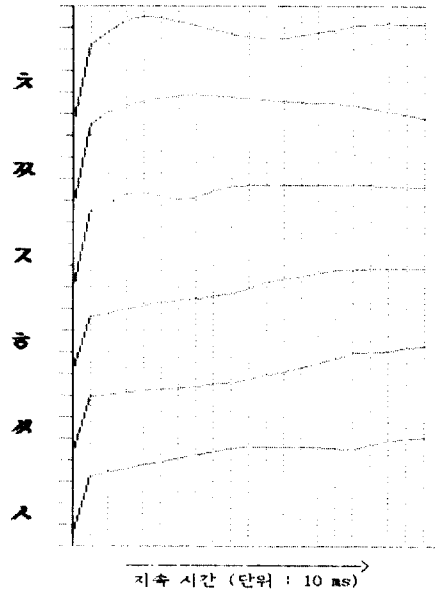


그림 6. power의 평균패턴 분포(후속모음이 /ㅈ/인 경우)

IV - 4 . 분산비의 분포

特徵量이 scalar量(單一特徵量)인 경우 class의 식별에 이용하는 특징량으로서 class 間의 분산이 크고 동시에 class 內의 분산이 작은 것이 요구된다.

이와 같은 특징의 평가량으로서 兩分散의 比가 이용된다. 분산비의 정의식을 다음에 나타낸다.

$$F = \frac{1 / (g-1) \cdot \sum_{k=1}^g (\mu_k - \mu)^2}{1 / (g(n-1)) \cdot \sum_{k=1}^g \sum_{i=1}^n (X_{ik} - \mu_k)^2}$$

여기서, X_{ik} : class k 특징량의 i번째의 관측치

($i = 1, 2, \dots, n; k = 1, 2, \dots, g$)

μ_k : class k 특징량의 평균치

μ : 전체 특징량의 평균치

그림 7에 始端에서부터 150ms 구간 (15 frame)의 시간 변화에 대하여 각 音素群別 (5 group)로 나누어 後續 母音이 /ㅈ/인 경우에 대한 分散比의 分布를 표시한 것이다.

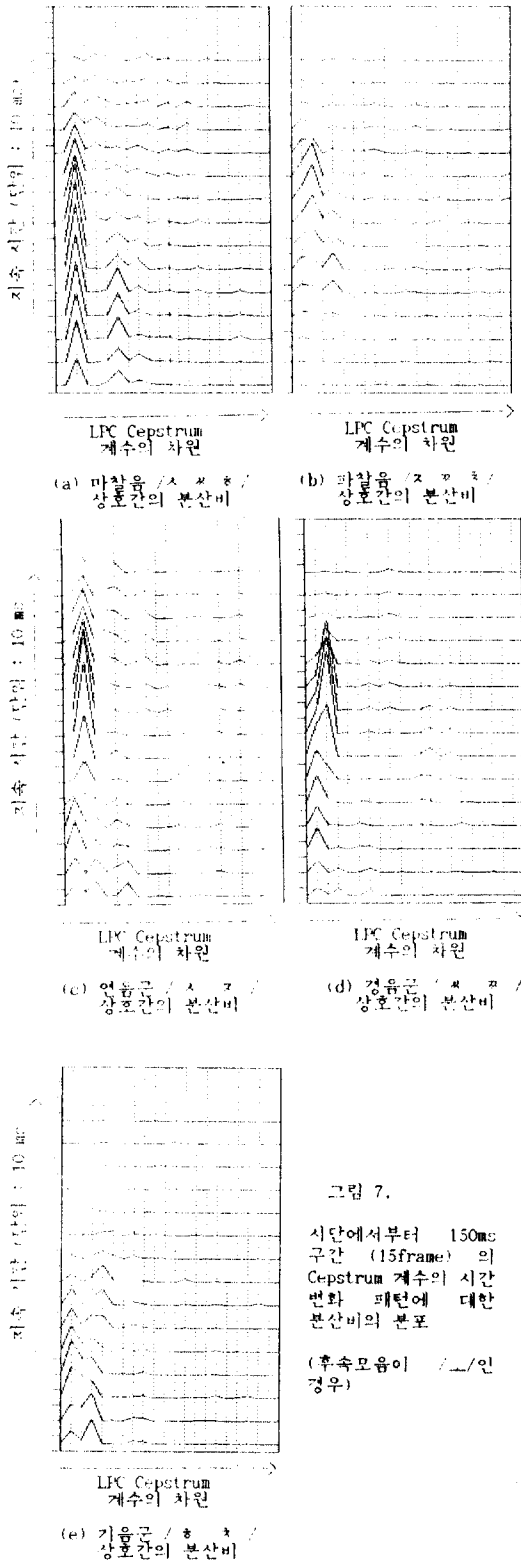


그림 7.

시단에서부터 150ms 구간 (15frame)의 Cepstrum 계수의 시간 변화 패턴에 대한 본산비의 분포

(후속모음이 /ㄷ, ㄹ/인 경우)

摩擦音 상호간의 본산비 (그림 7 a)에서는 Cepstrum 계수 C1의 본산비가 가장 크고 그다음에 C5이고, 破裂音 상호간의 본산비 (그림 7 b)에서는 C1, C3가, 軟音群 상호간의 본산비 (그림 7 c)에서는 C0, C1, C2, C5가, 硬音群 상호간의 본산비 (그림 7 d)에서는 C0, C1가, 氣音群 상호간의 본산비 (그림 7 e)에서는 C0, C2, C3의 본산비가 크게 나타나 전체적으로 C0 ~ C5의 Cepstrum 계수로 분리 가능함을 보여주고 있다.

摩擦音 상호간에서는 始端으로부터 제 14 frame까지의 分散比가 크게 나타나고, 破裂音 상호간에서는 제 4 frame까지는 본산비가 적어 식별에 대한 貢獻度가 적게 나타나지만 제 5 frame부터 제 11 frame까지의 分散比는 크게 나타나서 이 부분의 파라미터가 식별에 중요함을 알 수 있다.

또, 軟音群은 제 14 frame까지가, 硬音群은 제 11 frame까지가, 氣音群은 제 8 frame까지가 식별에 대한 貢獻度가 크게 나타나 전체적으로 14 frame 정도의 길이를 인식 파라미터로 할 경우 식별 가능함을 알 수 있다.

V. 결 론

摩擦音 및 破裂音을 지속시간, 평균패턴, 본산비를 이용하여 분석한 결과 다음과 같은 사실을 확인 할 수 있었다.

1) 持續時間의 分布로부터는

지속시간 정보가 마찰음에서는 軟音과 硬音의 分離에, 파찰음에서는 硬音, 軟音으로부터 氣音의 分離에 尤效하고,

Cepstrum 계수의 平均패턴 分布와 power의 平均패턴 分布로부터는 마찰음 파찰음의 分離가 가능함을 알 수 있었다.

2) 分散比의 分布로부터는

C0 ~ C5의 Cepstrum 계수로 분리 가능하고, 전체적으로 14 frame 정도의 길이를 認識 파라미터로 할 경우 摩擦音 및 破裂音의 識別이 가능함을 알 수 있었다.

3) 또, 分散比와 平均패턴으로부터는

6次元 정도의 Cepstrum 계수만으로도 摩擦音, 破裂音의 識別이 可能하다고 판단된다.

V. 참고 문헌

- [1]. 은종관 외 : "음소모델링을 이용한 한국어 대용량 단어 인식에 관한 연구" 음성통신 및 신호처리 WORKSHOP, 19-24 (Aug, 1989)
- [2]. 許 熊 : "국어 음운학" 정음사, 190-213, (1985)
- [3]. 우메다 : "한국어의 음성학적 연구", 형설출판사, 35-37 (1983)
- [4]. 정, 마키노, 키도 : "韓國語 破裂子音의 分析" 日本 音響 學會 1-3-3 (Oct, 1988)
- [5]. 정, 마키노, 키도 : "韓國語 語頭 破裂子音의 認識" 日本 音響 學會 1-2-21 (Mar, 1989)