

# TRACE : 상호작용 음성인식 Neural Network 모델

金 明 源  
(한국전자통신연구소 책임연구원)

## ■ 차 례 ■

- |                           |                      |
|---------------------------|----------------------|
| ① 개 요                     | 3.4 Context에 의한 음소인식 |
| ② TRACE model 의 구조        | ④ 단어 인식              |
| ③ 음소 인식에 영향을 미치는 요인       | 4.1 분리 단어 인식         |
| 3.1 어휘 효과(lexical effect) | 4.2 단어의 분리           |
| 3.2 Phonotactic 규칙 효과     | ⑤ 결 론                |
| 3.3 보완 관계와 Category인식     |                      |

## 요 약

본 논문은 음성인식을 위한 neural network의 하나인 TRACE model에 대하여 기술한다. TRACE model은 HEARSAY 음성이해 system의 blackboard 구조와 유사한 trace 구조를 사용하여 정보처리 unit들간의 자극성 내지 억제성 연결에 의한 상호작용의 결과로서 음성을 인식한다. TRACE model은 특징층, 음소층 그리고 단어층의 3층으로 구성되며 층간의 unit들이 상호작용할 뿐 아니라 동일층의 unit들이 상호경쟁함으로써 음성인식에 있어서의 context 효과, segmentation 및 잡음 등의 문제를 해결할 수 있다.

### ① 개 요

TRACE model은 음성인식에 있어서의 여러가지 문제점, 즉 1) 음성신호의 시간성, 2) context 효과, 3) 음성단위간 경계의 불확실성, 4) 잡음 등의 문제를 해결하기 위한 neural network model이다. TRACE에서 채택한 접근 방법은 HEARSAY[Lesse 75]음성이해 system에서 사용

된 blackboard 구조와 유사한 trace를 사용하여 처리 unit들간의 상호 자극성 내지 억제성 연결에 의한 상호작용의 결과로서 인식을 수행한다. 이 점에서 음성인식에 적용되는 다른 connectionist model 예컨대 TDNN(Time Delay Neural Network)[Waibe 89], MLP(Multi-Layer Perceptron)[Kamme 89] 및 INN(Integrated Neural Network)[Matsu 89]등과 그

구조 및 기능상에서 상이하다. TDNN, MLP 및 INN이 feed-forward의 network로서 하층 단계에서는 primitive한 feature들을, 그리고 상위 단계로 갈수록 좀 더 고차원적 추상적 feature들을 검출하여 최종적으로는 그들 추상적 feature들의 pattern으로써 음성신호를 분류(인식)한다. TRACE는 하층의 특징과 상층의 특징들이 상호작용할 뿐 아니라 동일층의 특징들이 상호경쟁함으로써 음성인식이 이루어질 수 있음을 제시한다.[McCle 86].

여기서는 TRACE model의 구조에 대해서 설명하고 다음에 TRACE model이 음성인식의 여러가지 문제점 및 현상을 설명할 수 있는 mechanism에 대하여 기술하고자 한다.

## 2) TRACE model의 구조

TRACE는 TRACE I과 TRACE II 두 개의 version이 있는데 이들은 각각 음성의 서로 다른 관점을 simulation하기 위해 개발된 것이다.

TRACE I : 실제 음성신호로부터 음성을 인식하는데 있어서 여러가지 문제점을 분석

TRACE II : 음성의 인식에 있어서 단어에 관한

지식이 미치는 영향 분석

TRACE model은 특징, 음소 및 단어의 3계층으로 이루어진 상호 활성화(interactive activation) neural network model로서 모든 unit들의 출력값이 동시에 계산된다는 의미에서 synchronous하다. 특징층에는 여러개의 특징 검출기(feature detector)들이 있는데 이들은 각각 음성신호의 어떤 특징을 검출한다. 각 특징 검출기는 시간상의 연속적 time slice마다 반복적으로 배열되어 있다. 음소층에는 각 음소에 해당하는 음소 검출기들이 있는데 한 음소에 대하여 복수의 동일한 음소 검출기가 매 3 time slice마다 반복적으로 배열되어 있고 하나의 검출기는 6 time slice를 cover한다. 단어층에는 각 단어에 대응하는 unit들이 있다. 음소층에서와 마찬가지로 단어 검출기도 매 3 time slice마다 반복적으로 배열되어 있고 한 검출기가 cover하는 영역은 단어의 길이에 해당한다.

그림 1은 TRACE 구조의 일부를 나타낸다. 여기서 각 사각형은 하나의 unit을, 사각형 내의 label은 그 unit에 대응하는 item을, 그리고 사각형의 수평경계는 각 unit이 cover하는 영역을 나타낸다. "Tea cup"의 음성신호에 대한(녹음이

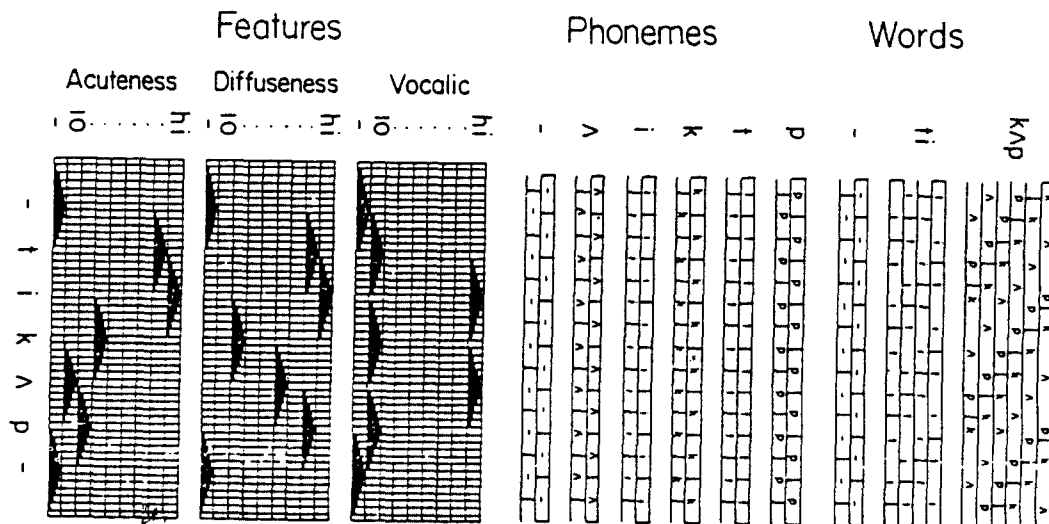


그림 1. TRACE의 구조도

전, 후에 있음) input 특징들이 해당 특징 unit에 검은 색으로 표시되어 있다.

TRACE에서 input은 연속적인 time slice의 특징 unit들에 순차적으로 즉 실제 음성신호가 input되는 것 같이 주어진다. 어느 한 순간 input은 특징층의 한 time slice의 특징 unit들에 주어진다.

TRACE model의 unit들의 전체 network를 "trace"라고 하는데 그 이유는 input 음성신호에 의한 각 unit의 반응 pattern이 3단계에 처리과정을 거쳐 결과된 input의 분석 data를 나타내기 때문이다. 여기서의 trace는 다른 trace와는 달리 unit들의 activation 값으로 되어 있고 또 이들

unit들은 지속적으로 상호작용한다. 이와 같은 상호작용이 TRACE model로 하여금 context 효과 뿐 아니라 어떤 information이 전체적으로 일관성있게 하나로 연결될 경우 비록 짧은 시간이라 할지라도 많은 양의 정보를 기억할 수 있다는 기억기능을 설명할 수 있게 한다.

TRACE에서의 정보처리는 unit들간의 자극성, 억제성 상호작용에 의하여 수행된다[McCle 81]. 서로다른 층의 unit들간에는 상호 자극성 연결이 그리고 동일층의 unit들간에는 상호 억제성 연결이 되어 있다. 그리고 모든 연결은 양쪽 방향이다.

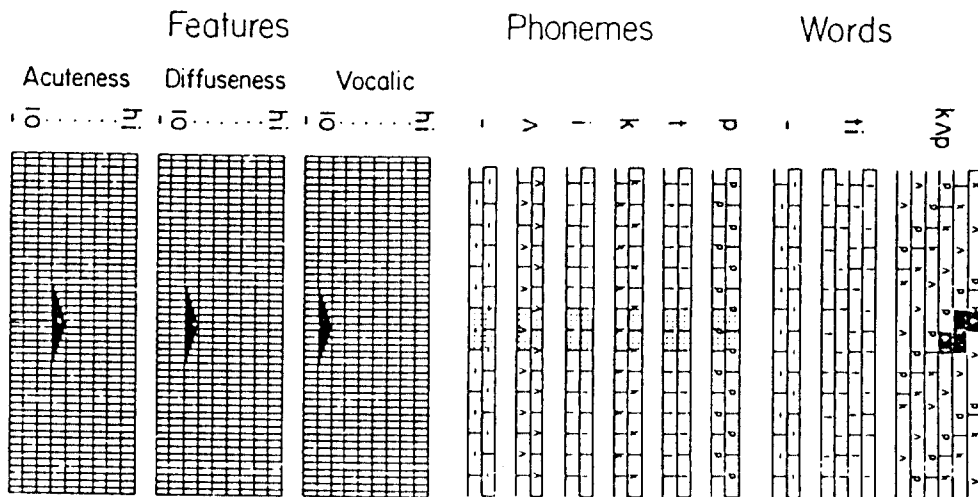


그림 2. 음소 /k/에 대한 unit들간의 연결

그림 2는 음소 /k/에 대하여 TRACE의 unit들간의 연결을 나타낸다. Time slice 24상에 중점을 둔 음소 /k/에 해당하는 unit가 단어층과 특징층의 unit들과 자극성 연결(검게 표시된 정도가 연결 강도를 나타냄)은 같은 층과 특징층의 unit들과 자극성 연결(회색 부분)을 가짐을 보여준다. 여기서 억제성 연결의 정도는 unit들간의 시간적 중복의 정도에 비례 한다. 유사한 연결법칙이 특징층 및 단어층의 unit에도 적용된다.

각 unit은 그것에 연결된 다른 unit들로 부터의 input을 합한다. 이 때 input은 문제의 unit의 output과 해당 unit들간의 연결 강도의 곱으로 나타내진다. 그 다음, unit activation은 다음과 같이 결정된다.

$$a'_i = a_i(1-\theta) + (\max - a_i)net'_i ; net'_i > 0일 경우$$

$$a'_i = a_i(1-\theta) + (a_i - \min)net'_i ; net'_i < 0일 경우$$

단,  $net_i = \sum w_{ij}a_j$ 이다. 여기서  $\theta$ 는 activation의 감소율을, 그리고  $\max$ ,  $\min$ 은 activation의 상한값과 하한값을 각각 나타낸다. 그리고 unit의 output은 다음과 같이 결정된다.

$a_i = a_i : a_i > 0$ 일 경우

0 :  $a_i < 0$ 일 경우

TRACE I의 구현 : TRACE I의 input은 15개의 parameter로(각 parameter는 그 값의 크기에 따라 8등급으로 나누어짐) 이루어지며, 한 남성 speaker가 발음한 음절을 5msec마다 추출한다. 사용된 음절은 자모음절로 자음(/p/, /t/, /k/)과 모음(/a/, /i/, /u/)의 결합으로 이루어진 것이다. Time slice는 100개로 전 system이 cover하는 음성신호의 시간길이는 500 msec이다. TRACE I에서는 단어층이 없고 unit들간의 연결강도는 context의 영향을 받는다. 한 음소로의 연결강도는 주위의 context의 음소들의 상대적 activation에 따라 결정된다.

표 1. TRACE II에서 사용된 음소별 특징 값

PHONEME FEATURE VALUES USED IN TRACE II

PHONEME	POW	VOC	DIF	ACU	CON	VOI	BUR
p	4	1	7	2	8	1	8
b	4	1	7	2	8	7	7
t	4	1	7	7	8	1	6
d	4	1	7	7	8	7	5
k	4	1	2	3	8	1	4
g	4	1	2	3	8	7	3
s	6	4	7	8	5	1	-
S	6	4	6	4	5	1	-
r	7	7	1	2	3	8	-
l	7	7	2	4	3	8	-
a	8	8	2	1	1	8	-
i	8	8	8	8	1	8	-
u	8	8	6	2	1	8	-
-	7	8	5	1	1	8	-

POW=power, VOC=vocaliness, DIF=diffuseness, ACU=acuteness, CON=consonantal, VOI=voicing, BUR=burst amplitude. Only the stops have values on this last dimension.

TRACE II의 구현 : TRACE II의 input은 실제 음성신호가 아니고 모의 음성신호이다. Simulation에 사용된 음소는 파열음 /b/, /p/, /d/, /t/, /g/, /k/, 마찰음 /s/, /S/, 유음 /l/, /r/ 그리고 모음 /a/, /i/, /u/, /-/이며 묵음(/-/로 표시됨)도 포함한다. 모의음성은 7개의 parameter(각 parameter는 8개의 등급이 있음)의 값으로 나타내진다. 표 1은 TRACE II에서 사용된 음소별 특징 parameter 값을 나타낸다.

Input 음성신호상의 각 음소에 대한 특징 parameter들은 11개의 time slice에 걸쳐 나타난다. 그 강도는 제 6 slice에서 최대이고 그 전후로는 값이 감소하도록 주어진다. 연속되는 음소에 대한 최대치가 6 slice 떨어져 있으므로 이들 음소간의 중복이 있게 된다. 음소층에는 15개의 음소에 해당하는 unit들이 있으며 이들 unit는 특징층의 11개 slice에 해당하는 unit들로부터 자극성 input을 받고 반면 이들 unit들에게 자극성 feedback 신호를 보낸다. TRACE II는 또 단어층에 200여개의 단어에 -묵음 포함- 해당하는 unit들을 가지고 있다.

TRACE에 있어서 unit들의 activation은 input 이 특징층을 통과함에 따라 그 값이 변하게 된다. 어떤 음소 혹은 단어인가의 최종 결정은 계산과정 중의 특정 시점에 하거나 어떤 unit activation이 선택 가능한 다른 unit의 activation의 주어진 상대적 값에 도달할 때 수행할 수 있다. 이 경우 어느 unit을 선택하는가는 Luce의 선택 법칙에 따른다. 즉

$$p(R_i) = \frac{S_i}{\sum S_i}$$

단,  $S_i = e^{ka_i}$

$a_i$ 는 unit i의 activation,  $p(R_i)$ 는 unit i를 결과로 선택할 확률을 나타냄.

### [3] 음소 인식에 영향을 미치는 요인

지금까지 TRACE model의 구조에 대해서 설명하였거니와 다음에는 특히 TRACE II를 사용하여 음소의 인식에 있어서의 단어 정보가 미치는 영향 및 단어 인식 mechanism에 대하여 기술한다.

음소의 인식에 영향을 미치는 요인으로서는 어휘효과(lexical effect), phonotactic규칙, parameter 간의 보완 관계와 category 인식 그리고 context 효과 등이 있는 바 여기서는 TRACE model이 갖는 병렬성과 상호작용 기능으로 이들 효과를 설명할 수 있는가를 기술하고자 한다.

(1) 어휘 효과(lexical effect)

어휘 효과는 음소 혹은 subword 단위의 인식에 있어서 단어에 대한 정보, 예컨대 그 단어가 있느냐 혹은 system의 어휘에 등록되어 있느냐

아니냐에 따라 부분적 인식이 영향을 받는 것을 말한다. TRACE의 simulation에서는 /b/와 /p/를 같은 정도로 activate하는 input 특징에 이어 /l/, /r/ 그리고 /g/에 대한 input 특징이 연속적으로 나타나는 input 음성신호를 사용하였다. 그림 3은 TRACE의 음소층과 단어층의 unit들의 activation 정도로 나타낸다. 단어층의 각 unit는 작은 사각형으로 표시되어 있으며 사각형의 위치는 해당 unit의 activation 정도를 나타낸다. X-축 방향의 “~” 표시는 snapshot이 취하여진 시점에서의 input의 processing점을 가리킨다. 이 그림은 첫 음소의 인식이 다음 음소들의 activation에 따라 점차 확연해짐을 보여준다. 제 4 panel이 예시하는 바와 같이 단어층에서 단어 “plug”이 경쟁에서 이김으로써 그 feedback 정보가 음소층의 음소 /p/로 하여금 음소 /b/를 능가하도록 한다.

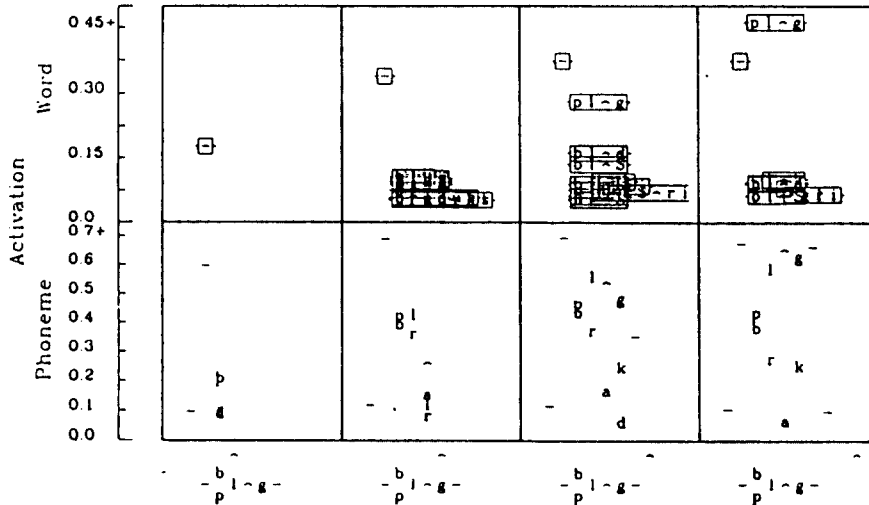


그림 3. 음소층과 단어층 unit들의 활성화 상태

(2) Phonotactic 규칙 효과

Phonotactic 규칙이란 영어에서 어떤 음소가 전후의 context에 따라 허용되거나 허용되지 않는 현상을 가리킨다. 예컨대 /t-i/에 대하여 /r/이 허용되는데 반하여 /l/은 허용되지

않는다. /s-i/의 context의 경우는 그 반대이다. 또 context /f-i/는 둘 다 허용되는데 context /v-i/의 경우 둘 다 허용되지 않는다. TRACE는 이와 같은 효과를 simulate할 수 있는데 그 이유는 모호한 input 자극이 여러

단어의 일부분을 activate하기 때문이다. 예컨대 sleep, sleet 등 단어의 처음 부분이 input /sɹi/에 의하여 activate된다. 여기서  $\uparrow$ 은 l음과 r음이 서로 반반씩 섞인 음을 말한다. 그러나 이들 단어들의 어느 것도 그 단어에 해당하는 input

이 주어졌을 때와 같은 완전한 반응을 보이지 않는다. 그러나 부분적으로 activate된 단어들이 공동으로 음소 /l/의 activation에 기여한다.

그림 4는 context /s-i/에서는 /l/과 /r/의 모호한 음소를 처리하는 과정을 나타낸다.

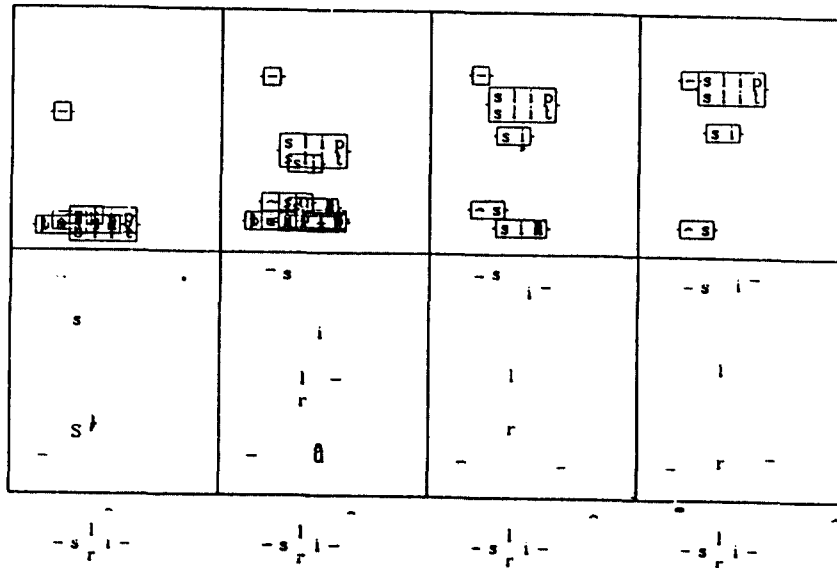


그림 4. Context /s-i/에서의 /l/과 /r/의 모호성을 처리하는 TRACE의 상태

### (3) 보완 관계와 category 인식

보완 관계(trading relation)는 음성인식에 있어서 두 개 이상의 특징 parameter들간에 상호 보완적 관계가 있는 것을 말한다. 예컨대 VOT(Voice Onset Time)과 F1OF(1st Formant Onset Frequency) parameter들 -실제 simulation에서는 voicing(VOI)과 burst amplitude(BUR) parameter가 이들 parameter로 대체됨- 간에 유사한 /k/ 음소와 /g/ 음소를 50대 50으로 선택하기 위해서 F1OF의 값이 높으면 높을수록 VOT의 값이 작아지는 관계가 있다. 즉 TRACE는 동일한 음소를 판별할 경우 input상의 차이를 경시하는 경향이 있다. 이와는 반대로, 상이한 음소를 구별할 경우에는 input상의 차이를 확대하는 경향이 있는데 이를 category

인식이라 하며, 이와 같은 현상은 음소층내의 unit들 간의 경쟁적 억제작용과 음소층에서 특징층으로의 feedback에 기인한다.

### (4) Context에 의한 음소인식

TRACE의 음소인식에 대한 조음효과(coarticulation)를 설명하는 데는 음소unit들의 activation이 특징층에서 음소층으로의 연결선의 강도를 조정함으로써 가능하다. 실제 TRACE I을 사용한 simulation에서 자모음절의 첫 자음을 인식하는데 연결 강도의 조정 결과 그렇지 않은 경우(인식율 79%)보다 향상된 인식율을(90%) 얻었다.

### [4] 단어 인식

TRACE model의 단어 인식 기능은 여러가지 면에서 기존의 COHORT model[Marsl 78]과 유사하다. 즉 단어인식에 있어서 top-down과 bottom-up 정보가 결합될 수 있는 것 외에 음성 인식의 여러가지 특성을 설명할 수 있다. TRACE model은 COHORT model의 장점을 가지고 있으며 동시에 COHORT model의 단점을 보완한 model이라 할 수 있다.

COHORT는 단어의 첫 음소에 해당하는 음성을 듣고 가능한 단어들의 집합을 -이를 cohort라 함- 결정한다. 다음 음소들이 처리되면서 match가 되지 않는 단어를 cohort로부터 제거해 나간다. 최종인식은 cohort에 남아 있는 단어가 유일하게 결정될 때이며 남아 있는 단어가 없을 경우 단어가 아니라는 결정을 내린다. 이와 같은 COHORT model은 몇가지의 단점이 있는 바 1) 단어의 처음 부분의 인식이 중요하므로 음성에 잡음이 있을 경우 인식에 결정적인 영향을 미칠 수 있고 cohort로부터 match되지 않은 단어를 제거하는 과정에서 일단 제거된 단어가 고려될 수 없으며 2) 단어의 처음과 끝이 분명히 분리되어 input되어야 한다는 제약이 있다. 여기서는 TRACE model이 COHORT model의 상술한 바의 단점을 해결할 수 있는 mechanism에 대하여 설명한다.

#### (1) 분리 단어 인식

TRACE에서는 COHORT의 첫번째 단점을 상호경쟁을 통하여 해결한다. 예컨대 "pleasant"와 "present"를 비교하면 "pleasant"의 input은 "present"보다는 "pleasant"를 더 많이 자극할 것이며 이는 단어 층내 unit들간의 경쟁적 억제 작용에 의하여 결국 "pleasant"가 다른 단어들을 독재하게 된다. 한편 "blacelet"의 경우는 이 단어가 없으므로 이것에 가장 잘 match되는 "bracelet"이 남는다. 단어의 처음 부분이 처리될 때까지는 "blame"이나 "blatant"와 같은 단어가 독재할 수 있으나 결국은 이들 단어들

보다 match가 더 잘 된 "bracket"가 최종적으로 독재하게 된다. 또 TRACE에서는 COHORT와 같은 계산이 진행되면서 맞지 않은 단어들이 제거되어가는데 그림 5는 이 과정을 보여준다.

단어 "product"(/prad kt/)를 처리하는 과정에서 음소가 순서대로 처리되어 가면서 후보 단어들이 탈락되어 간다. 즉 그림상의 곡선들은 "trot"(/trat/), "possible"(/pas b l/), "priest" (/prist/), "progress"(/prager^s/) 그리고 "produce(/pradus/)" 단어들이 차례로 탈락해 감을 보여준다. 여기서 TRACE가 COHORT와 다른 특징이 있는데 그것은 TRACE가 긴 단어보다는 짧은 단어를 더 선호한다는 것이다. 그 이유는 TRACE에서 단어층 unit들이 각 단어 unit의 cover하는 시간대의 중첩정도에 비례하여 상호 억제작용을 발휘하기 때문이다. 즉 긴 단어는 다른 단어들로 부터 짧은 단어보다 더 많은 억제신호를 받기 때문이다.

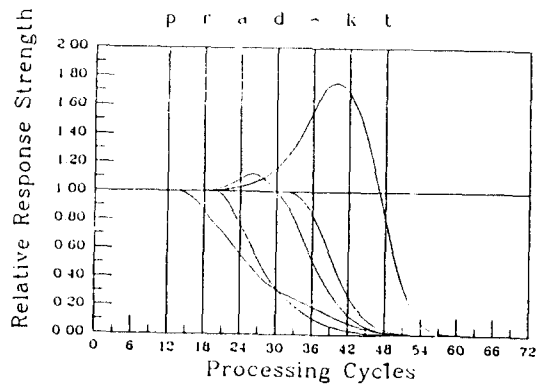


그림 5. 여러 단어들에 해당하는 unit의 반응강도 변화도

#### (2) 단어의 분리

음성인식에 있어서 단어의 분리 문제는 매우 중요하며 TRACE는 이 문제를 상호 작용 기능으로 해결한다. 즉 음소층에서 단어층으로의 자극성 연결과 단어층내의 경쟁적 억제기능이 그것이다. TRACE II를 사용한 simulation에서 임의의 단어들의 쌍을 만들어 두 단어를 단어사이에 어떤 구별가능한 것을 삽입하지 않은 것을

input으로 한다. 그림 6은 “possible target”( / pasˈbltɑrgˈt / )과 “pagusle target”( / pagˈsltɑrgˈt / )을 input하였을 때, 첫번째 경우 단어 “possible”의 끝이 어디인가가 다음 단어가 처리되기 시작한 초기에 분명해지는데 반하여 두번째의 경우는 다음 단어가 거의 다 처리될 때까지도 단어의 분리가 분명하지 못한 것을 보여준다. 이와 같은 현상은 단어 분리에 대한 어휘적 효과로서 TRACE와 같은 상호작용 mechanism이 인식에 있어서 의미있는 단위로의 분리 기능을 수행하는 것을 보여주는 것이다.

또 단어의 분리 문제에 있어서 TRACE는 짧은 두 개 단어로의 해석보다는 긴 한 단어로

해석하는 경향이 있다. 예컨대 “party”의 input에 대하여 “party”와 “par tea”의 2가지 해석이 가능한데 계산과정의 초기에는 “par”가 “party”를 독재하다가 시간이 갈수록 “party”가 독재하게 된다(그림 7).

그림 8은 연속 음성 / SiSˈtˈbaks / 를 처리하는 과정의 여러 시점에서의 TRACE의 상태를 나타낸다. 이 경우 문장 *she shut a box*가 최종적으로 독재함을 보여준다. 이는 TRACE model이 연속 음성으로부터 단어를 분리하는 데 있어 구문론적(syntactic), 의미론적(semantic) 제약 및 강세, 음절 규칙 등을 고려하는 계산 mechanism을 제시한다.

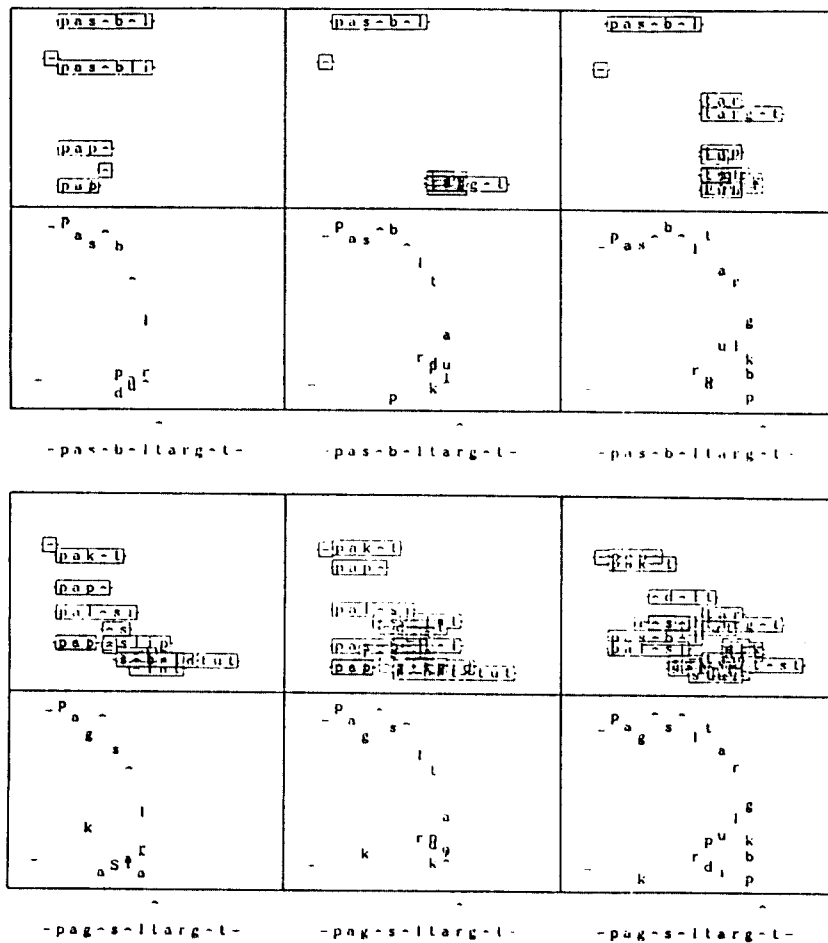


그림 6. possible target과 pagusle target을 처리하는 TRACE의 상태



- p	- p	- p	- p
a	a	a	a
r	r	r	r
u	u	u	u
b l a	u i d	p	
t	d	u	

-parti-      -parti-      -parti-      -parti-  
 그림 7. /parti/를 처리하는 과정의 TRACE

- s	- s	- s
i	i	i
s	s	s
l	l	l
b	b	b
a	a	a
o	u	u
	a	a
	u	u

-SiS-l-baks-      -SiS-l-baks-      -SiS-l-baks-

- s	- s	- s
i	i	i
s	s	s
l	l	l
b	b	b
a	a	a
k	k	k
s	s	s

-SiS-l-baks-      -SiS-l-baks-      -SiS-l-baks-

그림 8. 연속 음성 /SiSibaks/ (she shut s box)를 처리하는 TRACE의 상태

## 5 결 론

지금까지 TRACE가 음성인식에 관한 여러가지 현상을 simulation하는 것에 대하여 설명하였다. TRACE가 성공적인 이유는 두가지가 있다. 즉, 1) 대단위 병렬성과 상호작용, 2) trace 구조가 그것이다. TRACE에서는 음성신호를 처리하는데 있어서 음성 단위에 대한 가능한 가정이 나열되고 이들 사이의 상호 자극성 내지 억제성 연결을 통하여 하나의 일관된 가정이 다른 가정들을 독재하게 되며 이때의 독재하는 가정이 최종인식 결과가 된다. 음성인식 등과 같은 문제를 해결하는데 있어 문제의 구조를 적절히 잘 나타낼 수 있는 mechanism이 중요하다. 그러나 이것에 덧붙여 이에 상응하는 처리 system도 또한 중요하다. 음성이해 system인 HEARSAY가 그 구조는 적절하나 TRACE가 가지고 있는 것과 같은 대단위 병렬적 상호작용 기능을 결합함으로써 큰 성공을 거두지 못한 사실이 이를 입증하고 있다. TRACE model은 음성인식을 위한 다른 neural network model 예컨대, TDNN, MLP 및 INN 등과는 구조와 기능이 상이하다. 특히 다른 model들이 forward 연결만 되어 있어 음성정보가 한 방향으로만 처리됨으로써 인식에 있어서의 인식 단계간의 상호작용 측면을 설명하지 못한 데 반하여 TRACE model은 인식상의 상호작용 구조를 제시한다.

TRACE는 상술한 바의 장점이 있는 반면 몇가지의 보완되어야 할 결점이 있는 바, 그들은 다음과 같다.

1) 다수의 unit들과 연결선들이 특징층에서 음소층으로, 음소층에서 단어층으로서 연결형태의 반복적 복사로 인한 중복이 있다. 이러한 중복은 부분적 학습 문제를 야기한다.

2) TRACE model은 말하는 속도, 말하는 사람의 특성 및 accent 등의 전체적 parameter의 변화에 둔감하다.

3) TRACE model은 어떤 단어가 똑 같은 단어를 매우 짧은 시간 후에 다시 밀렸을 때, 그 단어의 인식에 영향을 미치는 것을 설명하지

못한다.

이와 같은 결점은 TRACE model의 구조를 수정, 보완함으로써 해결될 수 있을 것이다.

## 참 고 문 헌

- [Kamme 89] B.R. Kammerer and W.A. Kupper, "Design of Hierarchical Perceptron Structures and Their Application to the Task of Isolated-Word Recognition", Proc. of International Joint Conference on Neural Networks, 1989.
- [Lesse 75] V.R. Lesser, R.D. Fennell, L.D. Erman, and R.D. Reddy, "The HEARSAY II Speech understanding system", IEEE Trans. Acoustics, Speech, and Signal Processing, Vol. 23, Feb. 1975.
- [Mars 78] W.D. Marslen-Wilson and A. Welsh, "Processing Interactions and Lexical Access During Word Recognition in Continuous speech", Cognitive Psychology 10, pp. 29-63, 1978.
- [Matsu 89] T. Matsuoka et al., "Syllable Recognition Using Integrated Networks", Proc. of the International Joint Conference on Neural Networks, 1989.
- [McCle 81] J.L. McClelland and D.E. Rumelhart, "An Interactive Activation Model of Context Effects in Letter Perception : Part I. An Account of Basic Findings", Psychological Review 88, pp. 375-407, 1981.
- [McCle 86] J.L. McClelland and J.L. Elman, "The TRACE Model of Speech Perception", Cognitive Psychology 18, pp. 1-86, 1986.
- [Waibe 89] A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, and K.J. Lamg, "Phoneme Recognition Using Time-Delay Neural Networks", IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. 37, No.3, March, 1989.



金 明 源

---

저자약력

- 1972년 : 서울대학교 공과대학 응용수학 (학사)
- 1981년 : Univ. of Massachusetts (Amherst), computer science (석사)
- 1986년 : Univ. of Texas (Austin), computer science (박사)
- 1975년~1978년 : 한국과학기술연구소, 연구원
- 1982년~1985년 : Univ. of Texas (Austin), Institute for Computing and Computer Application, 연구원
- 1985년~1987년 : AT & T Bell Labs., 연구원
- 1987년~현재 : 한국전자통신연구소, 책임연구원