

技 術 解 說

# 인 간 — 컴 퓨 터 의 대 화

石 玟 秀\* · 宋 五 永\*\*

■ 차 례 ■

- I. 서 론
- II. 언어의 특성
- III. 단절 언어의 인식
- IV. 연속 언어의 인식
- V. 응 용
- VI. 결 론
- 참고문헌과 기타

## I. 서 론

인간-컴퓨터의 대화 방식도 컴퓨터의 발달 과정에 따라 변천되어 왔다. 초기의 기계적 Switch Machine Language 등 기계 중심의 대화 방식에서 Assembly Language, High Level Language (Fortran, Pascal, LISP 등)의 개발을 거쳐 점점 인간 중심의 대화 방법으로 발전되어 왔다. 그러나 이러한 High Level Language에 의한 대화 방식도 그 Language에 대한 숙련된 지식을 필요로 하므로 인간이 사용함에 있어 아직도 많은 제한성을 갖고 있다. 이러한 제한성을 극복하기 위한 수단으로서 컴퓨터가 직접 인간과 대화를 하도록 Natural Language Communication을 생각하게 되었다. 이러한 방식에는 많은 어려움이 있으나 VLSI발전과 Pattern Recognition, Artificial Intelligence 분야의 활발한 연구로 점차로 극복되어 머지 않은 장래에 인간과 대화하는 완전한 컴퓨터가 개발될 전망이다. 현재까지 개발되어 있는 컴퓨터도 간단한 언어 인식 능력을 갖고 있다. 그러나 이러한 컴퓨터의

언어 입력은 어떤 형식에 제한되어 있다. 이를테면 한번에 한명의 사람만이 컴퓨터와 대화를 할 수 있으며 16~300단어의 제한된 어휘만을 인식한다. 이러한 어휘는 각 단어마다 시간 간격을 두고 말해져야 한다. 이렇듯 언어 인식 시스템 연구의 중요한 목표는 기계에 대한 언어 입력의 기술 개발에 있다. Table 1은 각 시스템 고유의 난점에 따라서 분류된 여러 가지 다른 시스템들을 보여 준다. 여기서 “—”는 시스템에 따라서 다르기 때문에 기술되지 않았음을 뜻한다.

본 논문은, Natural Language에 의한 컴퓨터와의 대화방식의 설명을 위해, Section II에서는 언어의 특성과 이러한 특성을 추출하는 기법을 기술하며 Section III는 단절 언어 인식 시스템을 기술한다. Section IV에서는 연속 언어인식의 설명과 Section V에서는 이러한 인간-컴퓨터의 Natural Language에 의한 대화 방식이 실제로 현재 응용되고 있는 사례를 보여 준다.

## II. 언어의 특성

언어는 시간에 따라 변하는 복잡한 Pressure Wave

Table 1. 고유의 난점에 따른 언어인식 시스템의 분류

	언어형식	어휘크기	사용의 제한성	언 어	話者 언어형식	환 경
단절언어인식	단절된 언어	10~300	제한적	—	아주 협조적	—
연결언어의 제한인식	연결언어	30~500	제한적	제한된 명령어	아주 협조적	조용한 환경
언어의 제한이해	연결언어	100~2,000	무제한적	영어와 유사	협 조 적	—
제한적 Dictation Machine	연결언어	1,000~10,000	제한적	영어와 유사	아주 협조적	조용한 환경
언어 무제한 이해 Machine	연결언어	무 제한	무제한적	영 어	협 조 적	—
연결언어의 무제한 이해 Machine	연결언어	무 제한	무제한적	영 어	협 조 적	조용한 환경

\* 正會員 : 韓國科學技術院 電氣 · 電子工學科 副教授

\*\* 正會員 : 韓國科學技術院 電氣 · 電子工學科 碩士課程

이다. 그 과정은 발생된 단어들의 음소 나열에 대응하는 일련의 음향상태이다. 이러한 음향의 특징은 매우 가변적이기 때문에 언어 인식 시스템의 높은 정확도를 얻기는 매우 어려운 일이다. 음소 가변성의 중요한 요인들은 개개인의 Vocal Track의 체질적 차이와 한 사람이 말하는 동일 언어의 시간적 가변성 등이 있으며 또한 Vocal Track이 한 음소를 만들 때 話者が 말하려는 다음 음소에 따라서 준비를 하게 되는 특성이 있다. 예를 들어 'Cat'과 'Cam'에서 'a'라는 모음은 동일한 것일지라도 그 모음의 음향적 특징은 다르다. 똑같은 단어를 말하더라도 話者마다 발생 특징이 다르며 방언에 의한 가변성도 존재한다. 이러한 언어 특징의 다양한 가변성으로 인하여 언어 특징 추출은 쉽지 않으나 이것을 극복하기 위해 여러 가지 기술을 연구 개발하고 있다. 이러한 기술들 중에서 가장 널리 쓰이고 있는 것이 선형 예고 해석법이다. 이 방법의 기본적인 아이디어는 한 개의 언어 표본이 이미 말해진 언어 표본의 선형 조합으로써 근사화될 수 있다는 것이다. 실제의 언어 표본과의 차이의 제곱의 합을 극소화함으로써 예고시스템의 계수의 유일한 해의 집합을 결정한다. 이 방법은 기본적인 언어 특징인 Pitch, Formants, Spectra, 목청기능도 등을 추출하는데 이용된다. 그림 1은 이러한 해의 집합에 의해 언어발생 Model을 도식화한 것이다.

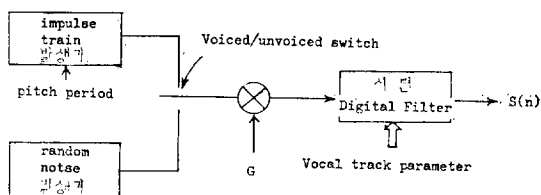


그림 1. 언어발생의 간략한 Block Diagram

Pressure Wave의 복잡한 Spectrum효과, Vocal Track, 성대 진동은 시변 Digital Filter에 의해 Modeling된다. 이 시변 Digital Filter의 Steady State System Function은 다음과 같다.

$$H(Z) = \frac{S(Z)}{U(Z)} = \frac{G}{1 - \sum a_k Z^{-k}}$$

이 Model의 시변 Parameter들은 Voiced/Unvoiced의 분별도, Voiced Speech의 Pitch Period, Gain Parameter (G), Digital Filter Coefficients {a<sub>k</sub>} 등이다. 이러한 Parameter는 언어 특징이라고 볼 수 있다. 이 Parameter에 대한 통계적인 처리를 행함으로써 발생된 언어를 인식할 수 있다. 다음 연속되는 2개의 Section에서는 단절 언어 인식과 연결 언어 인식에 대하여 기술할 것이다.

### III. 단절 언어의 인식(Isolated Word Recognition)

언어 인식의 복잡성은 언어가 연결되어 있느냐 아니면 한번에 한개의 단어만이 말해 지느냐에 의존적이다. 연결 언어에서는 한 단어의 시작과 끝의 위치를 결정하는 것이 어려우며 또한 그 단어들의 음향특성은 문맥에 따라서 매우 큰 가변성을 갖는다. 그러나 단절 언어 인식 시스템은 단어가 분리되어 인식되므로 이러한 어려움은 없다. 그림 2은 단절 언어 시스템의 Flow Chart를 보여 준다. 첫 단계에서는 발생된 음성 신호 즉 Analog Signal을 Digitize하고 분석하는 단계이다. 이러한 단계를 거친 음성 신호는 여러 가지 Parameter들로 표시되어야 한다. 이런 Parameter표시방법은 여러 사람에 의해 연구되었다. Itakura [1]는 선형 예고 부호화 계수(Linear Predictive Coding)을 사용하였고 Martin[2]은 Band Pass Filter에 의한 Hardware특징 추출법을 사용하였고 White [3]은 이외의 다른 Parameter표현법을 사용했을 때의 효과에 대해 연구하였다. 그 다음 단계에서는 단어의 시작과 끝을 찾아 낸다. 話者と Microphone이 동일할지라도 언어는 말해질 때마다 변하는 것은 물론이며 話者의 감정상태와 주변환경의 잡음에 따라 변하게 된다. 이러한 변화 요인들은 한마디의 언어가 발생되었을 때 그 언어의 진폭, Duration, SN비 (Signal-to-Noise ratio)의 변화를 야기하므로 Template Matching을 하기 전에 가변성을 극소화하기 위해 Normalization을 필요로 한다. Itakura는 잡음과 진폭의 Normalization을 수행하기 위해 전체의 발생 언어에 2차 Inverse Filter를 사용하였고 Martine은 Noise와 관련된 Room Noise, Breath Noise, 음절간의 Stop Gap 등과 같은 여러 가지 문제를 특성화하였다. 이러한 Noise는 단어의 시작과 끝을 부정확하게 찾아 내게 하는 요인므로 대부분의 Noise문제는 Breath Noise의 Spectra를 찾거나 단어들의 Mismatch된 음성을 버리는 등등의 기법을 사용함으로써 극복될 수 있다.

話者나 어휘가 바뀌었을 경우 시스템을 학습(Learning or Training)시켜서 Reference Pattern의 새로운 집합을 만들어 줘야 한다. Itakura와 White는 각 단어에 하나의 Reference Pattern을 이용하여 그 단어가 허용가능 이상의 Error를 갖을 때 그 단어를 새로운 Reference Pattern으로 저장하였다. Martine은 여러개의 표본으로부터 평균 Template를 발생시킴으로써 Reference Pattern의 추상화를 시도하였다.

미지의 단어 입력이 주어졌을 때 어떤 Reference-

의 선택 등과 같은 많은 설계의 선택 여지가 Performance에 따라서 존재한다. Itakura와 White는 Time Normalization에 대해 Dynamic Programming을 사용하였다. White는 Dynamic Programming에 근거한 비선형 의뢰어 선형 시간 분할 방법보다 좋은 결과를 보였다.

#### IV. 연속 언어의 인식

연속 언어 인식의 주된 목표는 문장과 같은 연속된 음성들을 정확히 인식하는 것이다. 그러한 시스템에서는 일련의 단어들은 단어 사이마다 Gap이 없이 자연스럽게 발생되어 지며 어휘들의 숫자도 훨씬 많아지며 언어 구조도 복잡해진다. 현재 연결 언어의 인식 분야는 계속 연구중에 있다. 이 중에서 대표적인 시스템은 "Systems Development Corporation (SDC)", "Bolt Bernak and Newman Inc.의 Hwim (Hear what I mean)", "Carnegie-Mellon University의 Hearsay 2, Harpy" 등이 있다. Table 2은 4가지 시스템의 Task Domain과 각각의 문법에 의해 허용되는 문장의 예를 보여 준다.

주어진 입력에 대한 단어 확인 과정은 쉽지 않으므로 입력데이터가 주어졌을 때 기대되는 단어가 존재할 확률을 측정하는 Mapping Technique이 필요하다. 이러한 역할을 해 주는 Mapper의 Performance를 세 가지 시스템에 대하여 비교한 것이 Table 3에 주어진다.

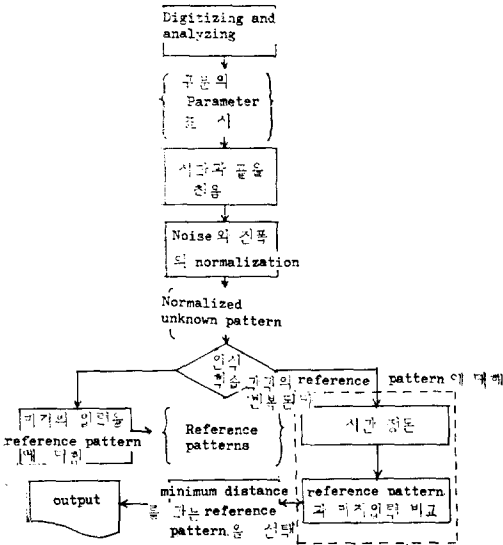


그림 2. 단절 언어인식 시스템의 Flow Chart

Pattern과 가장 유사성이 큰가를 결정하기 위해 모든 Reference Pattern과 비교된다. 이때 유사성은 미지의 입력과 Reference Pattern의 Minimum Distance 혹은 Maximum Correlation 등에 의해 판별된다. 이러한 단계는 "Matching and Classification"이라 일컬어진다. 이 단계에서는 Sampling Time Unit의 선택, Distance Metric의 선택, Parameter의 Differential Weighting, Time Normalization Function

Table 2. Task Domain와 문장에

Group	Task	Sample Sentence
SDC [4]	Facts About Ships	"How fact is the Theodore Roosevelt?"
BBN Hwim [5]	Travel Budget Management	"What is the plane fare to Ottawa?"
CMU Harpy [6]	Document Retrieval	"How many articles on psychology are there?"
Hearsay-II [7]	Document Retrieval	"How many articles on psychology are there?"

Table 3. Performance 통계자료

	확 인 결 정					
	SDC		BBN		CMU Hearsay II	
	ACCEPT	REJECT	ACCEPT	REJECT	ACCEPT	REJECT
옳은 단어	65 92%	6 8%	101 84%	19 16%	312 94%	20 6%
옳지 않은 단어	372 3%	11,253 97%	367 34%	713 66%	6,462 49%	6,591 51%
옳은 단어에 대한 단어가설설정	165		10		40	
음성의 유사성	RANDOM		5%		14%	

연속 언어 인식의 첫 단계는 음성 신호를 Filtering 해 주고 나서 Magnetic Tape에 기억시키는 것이다. 음성 신호에는 언어 잡음과 비언어 잡음 등이 섞여 있다.

비언어 잡음은 여러 다른 기술에 의해 제거된다. 그 다음은 다음과 같이 여러 단계를 거치게 된다.

**특징추출 :** 20KHz의 Sampling Rate로 Analog파형을 Digitize한 후 그 파형으로부터 음성 특징을 추출한다. 일반적으로 이러한 데이터는 주파수 영역 특징과 시간 영역 특징으로서 분류된다. 주파수 영역 특징은 음성신호의 각 주파수 영역에서 나타나는 Energy양을 결정하기 위한 Fourier 해석법 혹은 주파수 Filter를 이용한 Spectrum 측정장치로 부터 추출된다. 또한 선형 예고 부호 특성도 널리 사용된다. 시간 영역 특징은 시간의 주어진 주기와 각각의 주기에 대한 파형 특징과 같은 것을 포함한다.

**특징 분석 :** 특징 추출을 한 후에 여러 가지 다른 분석법을 사용하여 음성을 분석한다. 매 시간의 음성 Pattern은 다른 시간의 음성 Pattern들과 비교되어 가장 밀접한 Match를 이루는 음성 Pattern에 따라서 분류된다. 어떤 음성과 Correlation이 큰 음성 Pattern을 조사하는 분석법도 널리 사용되고 있다. 이렇게 찾아진 특징은 고주파 Energy, 비음발성시의 音流와 Vocal Track의 진동, 음성주기(Pitch Tracking), Vocal Track Resonants (Formants) 등을 포함한다.

**분할 :** 각 음소의 시작과 끝이 파형의 어느 부분인가를 결정하는 처리 과정이다. 각 음소는 이미 알려진 음소의 진본과 Match시킴으로써 분류된다. 이러한 처리를 하는 때 단계마다 일련의 음성 상태에 삽입, 제거, 치환 등을 해 줌으로써 Error를 줄여 나간다.

**음성 처리기의 Modelling :** 연속 언어 인식 시스템의 높은 정밀도를 얻기 위해 정확하게 음성 처리기를 Modelling하는 것이 중요하다. 이러한 일 중의 중요한 부분이 "Training"이다. 컴퓨터는 특별한 話者가 발성한 일련의 음성 상태의 적절한 Model과 음성 처리기의 결과를 비교한다. 그 다음 이 Model에 대한 Error의 종류와 주파수에 관하여 통계적인 처리를 행한다. 이 결과를 이용하여 추측된 Model을 일단 세운 후 주어져 있던 Model과 비교하면서 수정해 나간다.

**언어 Modelling :** 연속 언어 시스템은 허용 어휘수, 각 단어의 연결법, 각 단어의 사용 빈도 등을 규정해주는 Syntax에 대한 Software 언어 Model을 가져야 한다. Artificial Language는 어휘수가 적으며 문법적인 제한이 많으므로 Natural Language와 유사한 Compiler를 만들어야 한다.

언어 Modelling에 또 다른 중요한 것은 Semantics이다. 각 단어의 나열이 문법에 맞다고 해서 의미를 갖는 정보를 전달했다고 볼 수 없는 것이다. 불행하게도 이러한 Semantics에 관하여 알려진 것은 충분치 않다.

**언어 해독 :** 연속 언어 인식의 마지막 처리는 발생된 음성의 가장 확률이 높은 단어를 계산하는 언어 해독 과정이다. 확률은 2개의 중요한 정보를 적절히 이용함으로써 결정된다. 그 하나는 언어, 話者, 이미 결정된 Model들에 의해 정의된 음성 처리기 등과 같은 이미 알려진 특성으로 구성되어 있고 또 하나는 새로운 언어 신호 입력에 대한 음성 처리기의 출력이다. 언어 해독은 언어 신호의 특성이 알려진 단어와 그 단어들로 이루어진 구절과 Match되는 정도에 따라서 행해진다. 언어 해독의 중요한 부분은 음성 처리기 결과로서 나온 단어의 시작과 끝을 결정하는 것이다. 때로는 다른 구절이 똑같은 음성으로 처리되거나 그와는 역의 현상이 나타나기도 하며, 짧은 구절은 긴 구절에 포함됨으로써 잘못 분류되기도 하며 비음과 짧은 시간동안만 발생되는 자음 등이 크게 발생되거나 오랜 시간동안 발생되는 모음 등으로 인하여 잘못 인식되는 어려움이 있다. 또한 동음이의어에 의한 혼동도 있을 수가 있다. 그래서 정확히 단어를 해독하기 위한 여러가지 연구가 활발히 행해지고 있다.

## V. 응 용

미국의 Illinois주는 언어 인식 능력을 갖는 Credit Card System을 5여년의 연구 끝에 개발하였다. 이 시스템의 개발로 재정 절약 효과는 연간 \$220,000 상당에 이르는 것으로 보고되었다.

Caller가 중앙 주립 Credit Card Number를 물리면 컴퓨터와 직접 연결된다. Caller는 8단위 Credit Card Number를 전화기에 읽어 주면 컴퓨터는 이 번호를 반복하여 Caller에게 말해 주고 옳은 번호인가를 확인한다. Caller가 "yes"라고 대답하면 번호를 확인한 뒤 그 번호가 타당하면 Caller에게 Destination Number를 묻는다. 컴퓨터는 Caller가 말해 준 Number를 잘못 읽어 들었을 경우 컴퓨터가 옳은 번호인가를 물었을 때 Caller는 "no"라고 대답한다. 컴퓨터는 Caller에게 다시 Credit Card Number를 반복하도록 요청한다. Credit Card Number가 일단 확인이 되면 Caller가 Destination Number를 한 번에 한 단위씩 읽어주고 컴퓨터는 Caller에게 읽어 들인 번호를 반복해 준다. Destination Number가 옳게 읽혀졌을 경우 Call

은 끝나게 된다. 컴퓨터가 잘못 반복했을 경우 Caller가 "No"하면 컴퓨터는 "Sorry, Please Repeat the Digit"이라고 말한다. Caller는 다음 Digit을 말하기 전에 현재 잘못 읽혀진 Digit만을 다시 읽어 주면 된다. 이 언어 인식 시스템은 Drum에 기록된 남성의 목소리로 말하며 0~9 Digit과 40개 정도의 단어를 말할 수 있다.

미국의 Bell Laboratory는 인간과 컴퓨터가 Natural Language로 정보 전달을 할 수 있는 시스템 Simulation을 시도하였다. 그림 3는 이 시스템의 기능별 Block Diagram을 보여 준다. 이 시스템은 항공 예약 안내를 자동으로 해 주도록 되어 있다.

음성 처리기 (Acoustic Processor)와 Syntactic Parser는 서로 연관되어 있어 음성 처리기가 Spectrum 정보와 일치하도록 Parser에 의해 만들어진 가설의 단어를 시험해 본다. 모든 요소에 의해 공유된 2개의 Memory Units을 제외한 다른 나머지 부분은 語意 처리 과정 (Semantic Processing)에 속해 있다. 語意 처리기는 대화가 진행됨에 따라서 변화해 되는 World Model을 갖고 있으며 절대로 변화될 수 없는 Memory Module을 갖고 있다. World Model은 개념들의 집합

에 근거를 두고 있다. 그 개념들 중에는 목적지, 출발 날짜, 출발시각 등이 있다. 語意 처리기는 현재의 상태, 입력된 문장내의 단어들, 문장을 발생시키는 데 이용되는 State Diagram 등으로 부터 새로운 상태를 결정한다.

Memory Units은 두 종류의 정보를 저장하고 있다. 첫째는 「사실」에 관한 정보이다. 「사실」은 두가지 종류가 있다. 항공 계획은 Official Airline Guide의 한 부분으로 저장되어 있고 항공 안내에 있어서의 개념들 간의 상호 관계도 이 unit에 저장되어 있다. 이 시스템이 비행 경과 시간을 질문 받는다면 출발시각과 도착시각으로 부터 계산을 한다. 그렇게 하기 위해서는 각시의 표준시각을 기억하고 있어야 한다. 항공 안내를 할 때의 시작은 지방시각이기 때문이다.

둘째는 「절차」에 관한 정보이다. 「절차」는 입력과 World Model의 현 상태에서 새 정보를 도출하기 위해 기억된 「사실」을 사용하는 특별한 목적의 프로그램이다. 예를 들어 어떤 프로그램은 만세력에 대한 것이어야 한다. 그 프로그램은 어떤 날짜가 주어졌을 때 요일을 찾을 수 있다. 항공 예약 안내는 요일로 편성되어 있기 때문에 질문할 때 날짜로 말했을 경우 요일

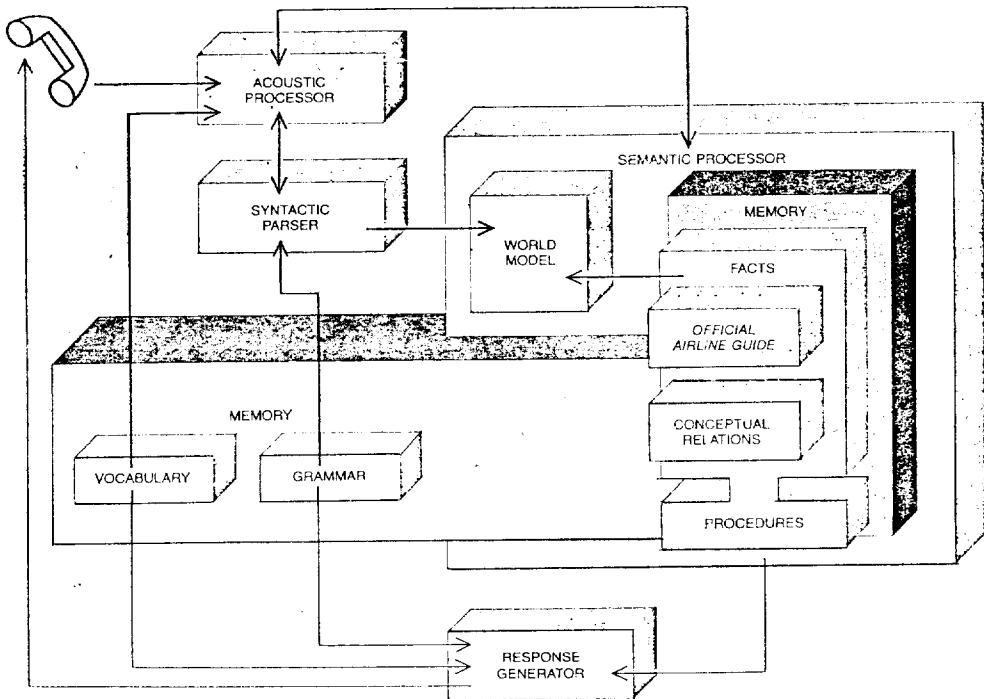


그림 3. Bell Lab System의 기능별 Block Diagram  
 굵은 선은 언어 인식에 관련된 정보의 흐름을 나타내며 가는 선은 응답의 발생 흐름을 나타낸다.

Machine: This is the Bell Laboratories flight information system. May I help you?  
 User: I want one first-class seat on flight number three one to Denver on Sunday.  
 Machine: I am sorry the flight is not available on Sunday.  
 User: I would like to leave on Saturday  
 Machine: Flight number three one leaves New York at nine forty-five A.M., arrives in Denver at one ten P.M.  
 User: What is the flight time?  
 Machine: The flight time is five hours and twenty-five minutes.  
 User: How many stops are there on the flight?  
 Machine: This flight makes one stop.  
 User: I want a nonstop flight.  
 Machine: I am sorry there are no nine A.M. nonstop flights.  
 User: I want to go at four P.M.  
 Machine: Flight number one six leaves New York at four fifty P.M. arrives in Denver at seven P.M.  
 User: What stops are there on the flight?  
 Machine: This flight makes no stops.  
 User: Is a Meal served on the flight?  
 Machine: Dinner is served.  
 User: How much is the fare?  
 Machine: The round-trip first-class fare is three hundred sixty dollars.  
 User: What plane is on the flight?  
 Machine: The aircraft is a Douglas DC ten.  
 User: I prefer the Boeing seven four seven.  
 Machine: The aircraft is a Douglas DC ten.  
 User: Please repeat the departure time.  
 Machine: The departure time is four fifty P.M.  
 User: I will take flight number six five.  
 Machine: I am sorry there is no flight six five to Denver.  
 User: Please repeat the flight number.  
 Machine: The flight number is one six.  
 User: I want flight one six.  
 Machine: Your reservation on flight number one six to Denver on Saturday is confirmed. Thank you.

그림 4. 항공 예약에 관한 전화 안내

로 환산하는 것이 필요하다.

시스템의 내부에서 사용자에게 대답을 해 줄때 시스템은 언어 부호 발생기를 동작시킨다. 語意 분석기는 언어 부호 발생기에게 World Model에서 전달될 개념들이 무엇인가를 알려 준다. 그러면 이 발생기는 Memory에서 문법과 어휘를 얻어서 개념들을 일련의 부호들로 바꾸어 준다. 언어 합성기는 이 부호들을 언어로 바꾸어 준다. 이 시스템의 대화에는 그림 4에 보여진다.

## VI. 결 론

연결 언어 인식의 응용은 시스템의 정교함, 생산 가

격, 설계자의 상상력 등에 의해 제한될 것이다. 그런 시스템이 최소한도로 갖추어야 할 것은 음성기억, 언어 처리, 話者에게 대담해 주는 Feedback, Error의 교정 등을 위한 부분들이다.

산업 사회에서 많은 종류의 컴퓨터, 자동조립, 생산 관리 등은 음성에 의해 제어될 수 있다. 음성에 의한 제어는 기계가 동작되는 장소와는 멀리 격리될 수도 있다. 언어 인식 Data의 처리 능력은 지역적으로 분산되거나 중앙으로 집중될 수도 있다. Data Base의 발달은 자동 언어 인식의 잠재성을 일깨워 줄 것임에 틀림 없다. 주식 시장에 대한 정보, 도서관 연구, 우편물 정리, 의학의 변천사, 대차대조표, 전화번호부, 각 실험실별 연구과제 등과 같은 다량의 정보를 처리할 수 있도록 대화식 컴퓨터가 이용될 것이며 언어학, 수학, 음악교습 등에도 이용될 수 있을 것이다.

그러면 언어 인식의 기술 발달은 어떤 방향으로 추진되어야 할까? 단기적으로는 언어 통신의 구조를 확실히 파악하는 것이 중요하다. 이것은 인간의 귀에 의해 행해지는 신호 해석에 관한 상세한 연구와 소리기호와 실질적인 소리와의 관계에 관한 보다 깊은 연구를 포함한다. 장기적으로는 여러 분야의 조사와 연구를 통하여 일반적인 이론을 세워야 한다. 또한 인간이 언어를 배우는 과정을 이해하는 것이 필요하다. 현재의 언어 인식 시스템은 「trained」이긴 하지만 이는 경험을 통해서 변화될 수는 없는 것이다. 이러한 적응 능력의 결핍은 언어 인식 시스템의 커다란 단점이며 난점이다. 언어 인식 시스템의 가장 좋은 설계 방법은 컴퓨터에 Natural Language와 같은 상세한 묘사 능력을 프로그램하는 것이 아니라 컴퓨터에, 언어를 배우는데 있어 필요한 능력의 기본 요소를 넣어 주는 것이다. 이러한 설계 방법이 Natural Language에 의한 통신에 얼마나 접근할 수 있을런지를 예측하기는 어렵다. 그러나 연구의 진척속도가 느리던 빠르던 이 목표는 끊임없이 추구될 것이다.

## 참고 문헌과 기타

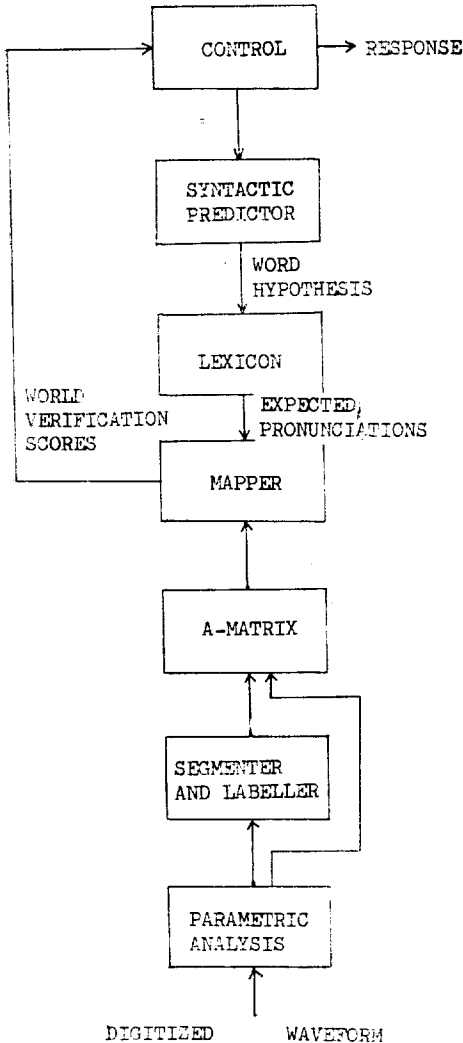
- (1) F. Itakura; "Minimum Prediction Residual Principle Applied to Speech Recognition", IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-23, pp. 67-72, Feb. 1975
- (2) T.B. Martin; "Applications of Limited Vocabulary Recognition systems", Invited Papers of the IEEE Symp. New York: Academic Press, 1975

[3] G.M. White and R.B. Neely; "Speech Recognition Experiments with Linear Prediction, Bandpass Filtering, and Dynamic Programming", in Proc. 2nd USA-JAPAN Computer Conf, Tokyo, Japan, Aug. 1975.

[4] SDC (System Development Corporation)

\*headed by B. Ritea

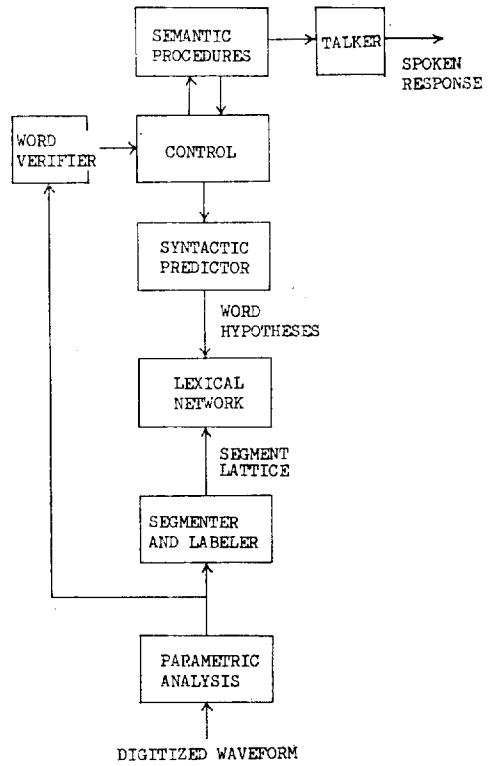
\*SDC System organization



[5] BBN (Bolt Bernak and Newman Inc.)의 Hwim (Hear what I mean)

\*Headed by W.A. Woods

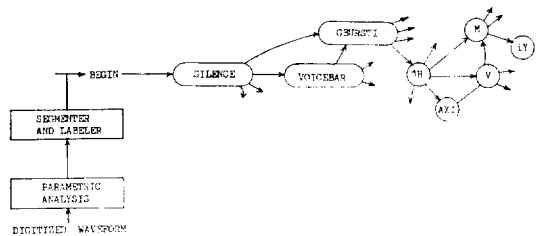
\*BBN Hwim system organization



[6] CMU (Carnegie-Mellon University)의 Hy

\*Headed by D.R. Reddy

\*CMU Harpy system organization



● "Give me..."에 대한 state transition network  
 ● 각 node는 선형예고 spectrum pattern에 따라 이름을 붙임.

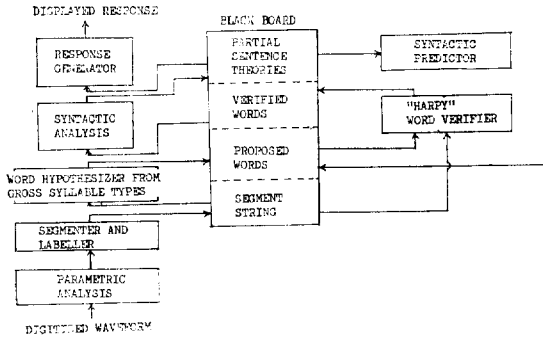
[7] CMU의 Hearsay-II

\*Headed by D.R. Reddy

\*CMU Hearsay-II system organization

[8] Aho, A.V., and Ullman, J.D; Theory of Parsing, Translation and Compiling (Prentice-Hall Englewood cliffs, NJ). 1972.

[9] Chomsky, N. and Halle, M; The sound Pattern of English (Harper and Row, New York)



1968.

[10] Reddy, D.R; Speech Recognition: Invited Papeors Presented at the 1974 IEEE Symposium (Academic, New York) 1975.  
 [11] Reddy D.R., et al; Speech understanding systems Final Report (Computer Science Department, CMU) 1977.

[12] Ritea, B; "Automatic Speech Understanding Systems," Proc. of the 11th IEEE Computer Society Conf. Washington, DC, pp. 319-322 1975.  
 [13] Woods, W., Bates, M., Brown, G., Bruce, B., Cook, C., Klovstad, J., Makhoul, J., Nash-Webber, B., Schwartz, R., Wolf, J., and Zue, V; "Speech Understanding Systems: Final Technical Progress Report," BBN Inc. Rep. No. 3438, Cambridge, MA 1976.  
 [14] Baker, J; "The Dragon Syntem-An Overview," IEEE Trans, Acoust. Speech Signal Process. ASSP-23, 24-29 1975.  
 [15] On Human Cammunication: A Review, A Survey and a Criticism. Colin Cherry. The MIT Press 1966.  
 [16] Trends in Speech Recognition, Edited by W. A. Lea. Prentice-Hall 1980.

### 電氣用語集 發刊案内

今般 當學會에서는 電氣工學을 研究하는 電氣人 및 學徒들의 學術研究에 寄與하도록 하는 한편 電氣業界 및 關聯分野의 現場에서 從事하는 實務者들에게 없어서는 안될 電氣用語集(18,000單語)을 發刊하여 低廉한 價格으로 보급하고 있아오니 많이 活用해 주시기 바랍니다.

會員 : 5,000원    非會員 : 5,500원    連絡處 : 267-0213 · 260-2253

社團法人 大韓電氣學會