

규칙기반과 신경망 모델을 결합한 한국어 글자-음소 변환 시스템 개발에 관한 연구¹⁾

김 세 훈, 이 주 현

한국 외국어대학교 경영정보 대학원
응용전산학과

서울시 동대문구 이문동 270번지

요 약

본 연구는 한국어 음성합성 시스템에서 한글 텍스트를 음소로 변환 시키는 규칙기반과 신경망을 결합한 한글-음소 변환 시스템을 제안하고 이를 위해 시스템 모델을 설계하고 시스템의 각 구성요소들을 설명하며 한국어 음운 변동 규칙중 설측음화 데이터와 설측음화에 상충되는 데이터를 사용하여 시스템을 실험하고 제안된 모형의 타당성을 분석한다.

1. 서론

음성합성 시스템은 입력으로서 임의의 텍스트를 받아 들어서 실시간 합성된 음성을 생성해낸다[10]. 음성합성 시스템은 언어적 정보에서 음성 파라메터를 추출하는 언어적 처리과정과 언어적 처리과정에서 추출된 음성 파라메터를 이용해서 실제적으로 음성을 합성하는 음성합성 처리과정으로 나눌수있다.

한국어 음성합성 시스템을 구현하기 위해서는 음성합성을 처리하는 한국어 음성합성기가 개발되어야 할 뿐만 아니라 언어적 처리과정 중 입력된 한글 텍스트를 정확한 음소로 변환시켜주는 연구가 필요하다.

입력된 텍스트를 정확한 음소로 변환시키는 방법에는 기존의 규칙기반 접근방법과 Sejnowski[14]의 신경망 모델 접근 방법 두가지가 있다.

기존의 규칙기반 접근방법은 각각의 규칙을 손으로 직접 기술해야 함으로 노동집약적이나 신경망 모델 접근 방법은 매우 복잡한 문제에 대하여 빠르고 좋은 해답을 찾을수 있다. 입력된 한글 텍스트를 정확한 음소로 변환시키는 것이 음성합성 시스템의 목적이고 규칙기반 접근방법과 신경망 모델 접근 방법이 서로 경쟁하는 것이 아니라 오히려 각각의 분야의 장점을 살려서 상호 보완적인 관계를 지켜 나가는 것이 바람직

본 연구는 한국 외국어 대학교 학술지원에 의한것임.

(이주현: "컴퓨터 보조교육 시스템을 위한 한글음성 재생 SW의 연구 개발")

하다[2].

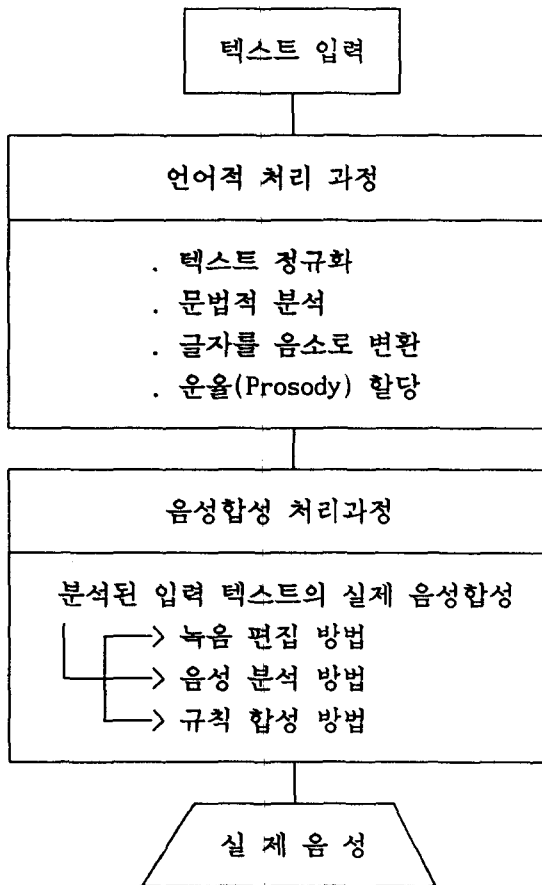
한국어 음성합성 시스템으로는 규칙 기반적 접근으로 개발한 시스템 공학센터의 텍스트-음성 변환 시스템[4]과 한국 전자통신 연구소의 '글소리' 시스템[5]이 있으며 신경망 모델접근 방법으로는 전자통신 연구소의 신경망을 이용한 한글단어-발음기호 변환 시스템이 있다.

본 연구는 한국어 음성합성 시스템 중에서 입력된 한글 텍스트를 정확한 음소로 변환시키는 규칙기반과 신경망 모델을 혼합한 글자 - 음소 변환에 대해 연구하였다.

2. 규칙 기반 음성합성

음성합성 시스템이란 데이터로서 저장되어 있는 언어적 정보 또는 텍스트를 음성으로 변환시켜 주는 것을 말하며 언어적 처리과정과 음성합성 처리과정과 같이 두개의 주요한 구성요소로 이루어져 있다[11][12].

음성합성 시스템의 구성은 [그림 2-1]과 같다.



[그림 2-1] 음성합성시스템 흐름도

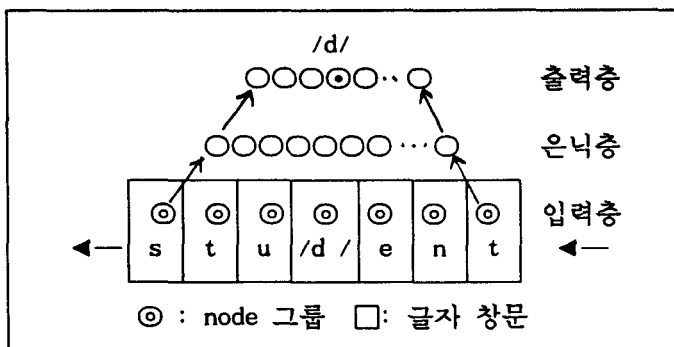
언어적 처리과정은 텍스트 정규화 단계, 글자-음소 변환 단계, 운율 할당 단계,

음성학 규칙 단계등으로 구분할수 있다. 첫번째는 텍스트 정규화 단계로서 실제 텍스트에서 나타나는 숫자, 생략어 그리고 그래픽 형식으로 표시된 정보등과 같이 상징적인 의미를 갖는 것을 표준 텍스트로 풀어 쓰거나 변환시켜 준다. 두번째는 글자-음소 변환(letter to phoneme) 단계로서 단어의 발음을 결정 하여 글자를 음소로 변환시킨다. 단어를 정확하게 발음하기 위해서 시스템은 각 글자의 나열을 그에 상응하는 음소의 나열로 상호 연결을 시켜 주어야 한다. 단어는 발음하고 강세 할당하는 데 있어서 규칙적인 패턴이 존재한다. 그렇지 않은 단어들에 대해서는 사전(dictionary)이 필요하게 된다. 사전은 문법적이거나 문맥 의존적인 단어와 발음법칙에 예외적인 단어를 처리하며 발음 규칙은 그 이외의 단어를 처리한다. 세번째는 운율(Prosody) 단계로서 리듬, 억양과 단어의 강조등과 같은 운율 파라메타를 텍스트에서 추출한다. 네번째 단계는 음성학 규칙(phonetic rule)으로서 음소에 길이(duration)를 할당한다.

음성합성 처리과정은 인간의 음성을 합성하는 방법으로서 녹음 편집 방법, 파라메타 편집방법, 규칙합성 방법등이 사용된다.

3. 신경망 모델을 이용한 음성합성

Sejnowski가 제안한 NETtalk[14]은 텍스트의 각 글자(letter)를 그에 해당하는 음소로 변환시키는 3층의 계층적 신경망 모델이다.



[그림 3-1] NETtalk 구조

입력층에는 7개의 node 그룹 또는 글자 창문이 있고 은닉층과 출력층에는 각각 하나의 node 그룹이 있다. 입력층의 각 node 그룹은 입력 텍스트의 한 글자를 나타낸다. 그러므로 입력층에는 항상 7개의 글자 나열이 있게 된다.

7개의 글자 창문중 네번째 글자, 즉 가운데 글자에 대한 정확한 음소가 이 네트워크에서의 출력값이 된다. 나머지 6글자는 가운데 글자의 음소결정에 부분적으로 영향을 주부는 글자들이다.

실험은 창문상에서 글자를 하나씩 움직여 가며 이루어 진다. 각각의 단계에 있어서 네트워크는 음소를 계산하고 각 단어에 대해 계산되어 나온 발음과 정확한 음소와의 차를 Rumelhart[8]등에 의해 소개된 역전파 학습알고리즘을 사용하여 출력층에서 입력층으로 역전파하여 연결강도를 조절해 나아간다.

입력층 node 갯수는 203개이며 은닉층의 node는 80여개로 구성되며 출력층은 26개의 발음기호를 나타내는 node로 구성된다[9][15].

NETalk은 웹스터 사전에서 1,000 단어를 선택해서 실험한 결과, 은닉층이 없을 경우 82%, 120개의 은닉층 node가 있을 경우 98%의 성능을 보였으며[14] 발음 장치인 DECTalk과 연결되어 문장을 소리내어 읽는 시스템으로 구성되기도 하였다.

4. 한국어 음운 변동과 연구 범위

음운변이는 그 변동의 성질에 따라 귀착, 동화, 이화, 축약, 생략, 첨가등으로 구분할수 있다[7].

음운변동에 있어서 가장 흔한것은 동화 작용이다. 조음위치가 다르고 조음방법이 다른 소리를 연달아 발음하기 보다는 조음위치나 방법이 같은 소리를 이어 발음하는 것이 쉽다. 따라서 음운변화의 대부분은 동화작용이고 동화작용은 어느 언어에서나 발견된다. 그러므로 본 연구에서는 동화작용 중 설측음화를 실험 대상으로 삼았다.

설측	법칙	'ㄴ'은 'ㄹ'의 앞이나 뒤에서 'ㄹ'로 발음한다.
음화	예	난로 -> 날로, 신라 -> 실라, 천리 -> 철리

[표 4-1] 자음접변 규칙 및 예

그러나 설측음화의 경우 'ㄴ'은 'ㄹ'의 앞이나 뒤에서 'ㄹ'로 발음하여야 하나 그 예외적인 경우가 존재하게 된다. 다음과 같은 단어의 경우는 'ㄹ'을 'ㄴ'으로 발음한다.

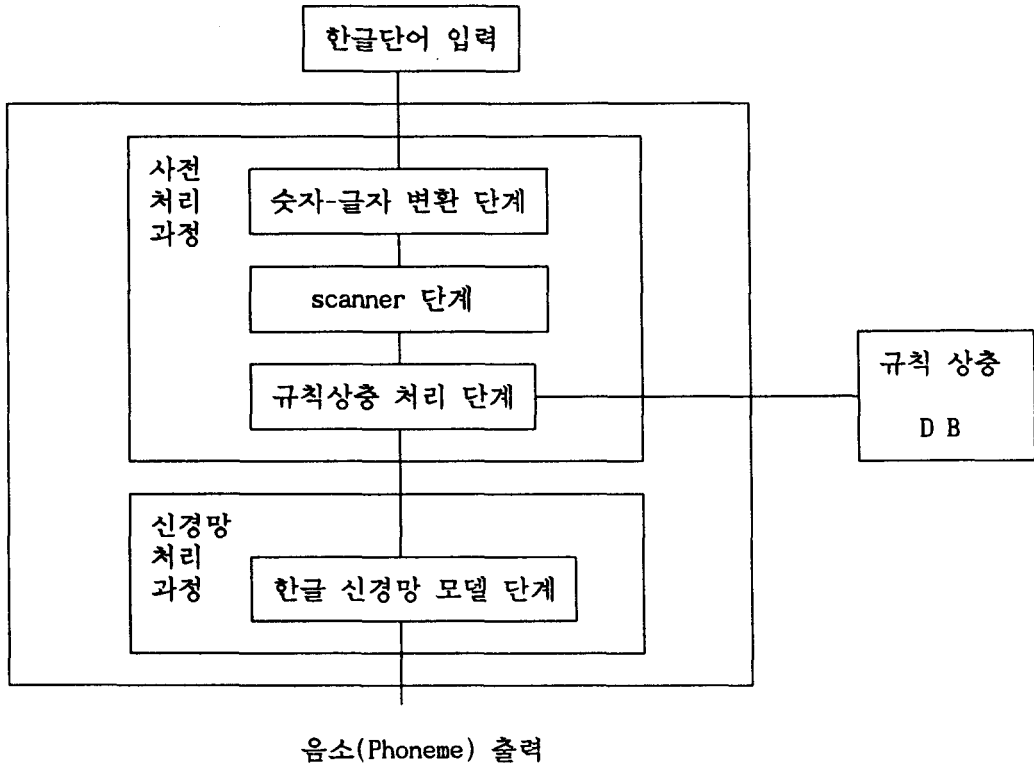
> 의견란 -> 의견난, 임진란 -> 임진난, 생산량 -> 생산냥 [6]

5. 한국어 글자 - 음소변환 시스템

실제 언어생활에 있어서 각 음운은 단독으로 발음되지 않는다. 음절이나 단어를 형성할때 음절과 음절이 연이어 소리 나거나 단어와 단어가 연이어 소리 날때, 여러가지 다른소리로 발음된다. 한글의 경우 규칙 기반으로 음운변동 규칙을 적용하였을때는 문제점이 있다. 그 이유는 음운변동 규칙의 적용순서가 다르기 때문이다. 일례로 '웃이'의 경우 실제발음은 연음법칙의 적용으로 인해 '오시'로 발음되나 '웃안'의 경우 실제발음의 연음법칙 적용 이전에 말음법칙이 적용되어 '온안'로 바뀌고 다시 연음법칙이 적용되어 '오산'이 아닌 '오단'으로 실제 발음된다. 그러므로 신경망에 이와 같은 '웃이', '웃안'을 '오시', '오단'으로 학습 시킴으로써 신경망을 각 음운변동 규칙 및 그 적용순서까지 학습시키는 결과를 낳게 된다. 그러나 신경망 모델은 학습시킬 규칙에 해당하는 단어와 상충되는 예외사항적 단어를 같이 학습시키게 되면 신경망이 제대로 학습되지 않기 때문에 이와 같은 경우는 규칙기반적 접근이 타당하다. 그리고 한글의 모아쓰기 특성인 초성, 중성, 종성의 위치 정보와 숫자-글자 변환은 신경망 이용보다는 규칙기반 접근 방법이 보다 유용하게 된다.

한국어 글자 - 음소변환 시스템은 크게 사전처리(Preprocessing) 과정과 한글 신경망 모델 과정으로 이루어진다. 사전처리 과정은 다시 스캐너(scanner)단계, 숫자 - 글자 변환 단계 그리고 규칙 상충 처리 단계로 구성된다.

본 한국어 글자 - 음소변환 시스템 구성은 [그림 5-1]과 같이 나타낼 수 있다.



[그림 5-1] 한국어 글자 - 음소 변환 시스템 흐름도

5-1. 숫자 - 글자 변환 단계

한글 텍스트에는 한글뿐만 아니라 아라비아 숫자가 자주 등장하게 된다. 예를들어 "1991년"의 경우 "일 구 구 일"이 아닌 "천 구백 구십 일"으로 읽게된다. 이러한 아라비아 숫자를 한글 신경망 모델에 입력시키기전에 한글로 변환시켜 신경망에 적합한 입력으로 변환 시켜준다.

5-2. scanner 단계

입력으로 들어온 단어를 먼저 각 음절에 대해 초성, 중성, 종성으로 분류하며 그 결과를 한글 신경망 모델로 보낸다. 한국어의 발음변동은 주로 중성과 초성, 즉 연이어 나오는 자음들 사이에서 발생하거나 또는 인접한 모음의 종류에 따라서 결정되므로 이와 같은 분류는 발음변동이 일어날수 있는 위치에 대한 정보를 제공 한다.

5-3. 규칙 상충 처리 사전 단계

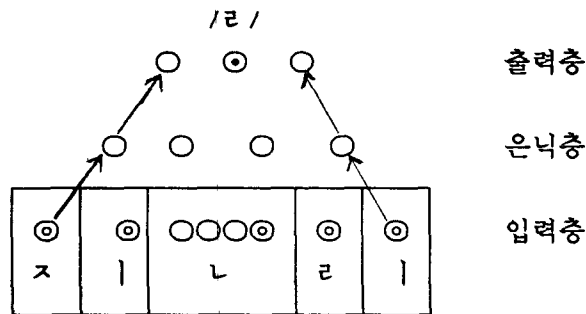
규칙 상충 처리 사전은 음운의 변화 규칙과 서로 상충되는 예외자료를 저장시키는 장소로 사용된다.

설측음화의 경우 'ㄴ'은 'ㄹ'의 앞이나 뒤에서 'ㄹ'로 발음한다. 그러나 '의견란'의 경우 '의견난'으로 발음한다. 이와같이 음운변동 규칙과 상충되는 예외적인 단어들을 저장하게 된다. 따라서 입력된 한글 단어들은 먼저 규칙 상충 처리 사전 조사를 거쳐 예외적 발음을 처리한 후 다음단계로 넘어가게 된다.

5-4. 한글 신경망 모델단계

한글 신경망 모델은 입력을 한글로 하며, 출력을 한국어 음운변동 규칙에 맞도록 변경된 음소로 출력하게 함으로써 구현될수 있다.

본 연구에서 제안한 한글 신경망 모델의 구조를 그림으로 표시하면 [그림 5-2]와 같다.



[그림 5-2] 한글 신경망 모델 모델

Sejnowski는 영어의 단어에 있어서 하나의 알파벳 발음은 그 알파벳 좌우로 3개씩, 모두 6개의 알파벳에 의해 영향을 받는다는 생각에 NETtalk의 입력층을 7개 그룹으로 구성[14]하였으며 한국 전자통신 연구소[3]에서 실험한 신경망모델을 이용한 한글 단어 발음기호 변환 시스템에서도 같은 구조로 실험하였다.

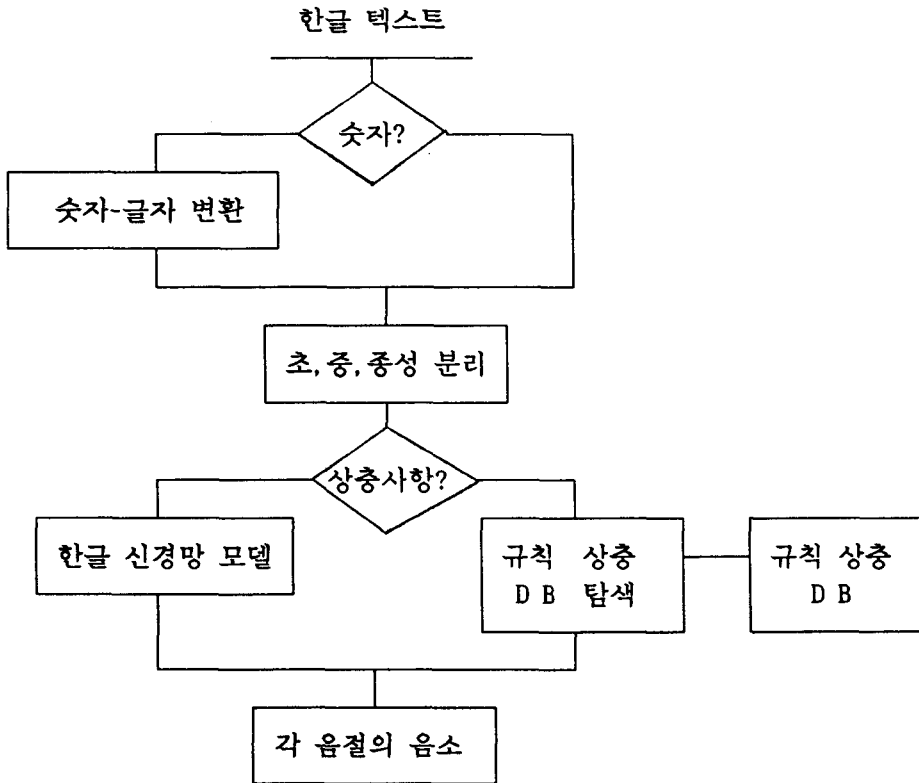
한글의 음운변동을 관찰해 보면 종성과 연이어 나오는 초성 또는 대립적인 종성의 영향에 의해 일어남으로 한글 신경망 모델 구조는 5개의 입력층이 적합하다고 생각하여 입력층을 5개의 node 그룹으로 구성하며 각각의 node 그룹에 3개의 node를 추가시켜서 한글 신경망 모델의 전 단계인 scanner단계의 정보 즉, 각각의 음운에 대한 초성, 중성, 종성에 대한 정보를 처리하도록 한다.

위와 같은 구조를 갖도록 한글 신경망 모델을 구성함으로써 음운변화가 주로 종성과 연이어 나오는 초성사이에 발생하는 한국어의 특징을 한글 신경망 모델에 심을수 있다.

6. 한국어 글자 - 음소 시스템 구현 및 결과

본연구에서 제안한 시스템의 각 구성 요소들의 구현 방법과 시스템의 실험 결과에 관해서 기술하였다.

전체적인 한국어 글자 - 음소 변환 시스템 흐름도는 [그림 6-1]과 같다.



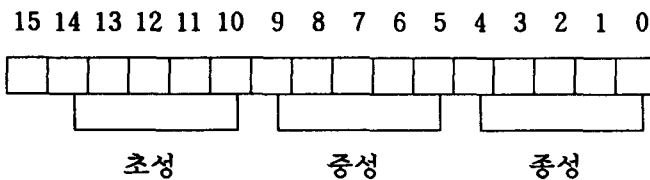
[그림 6-1] 한국어 글자 - 음소 변환 시스템 흐름도

6-1 전처리(Preprocessor) 단계 구현

1) scanner 처리단계

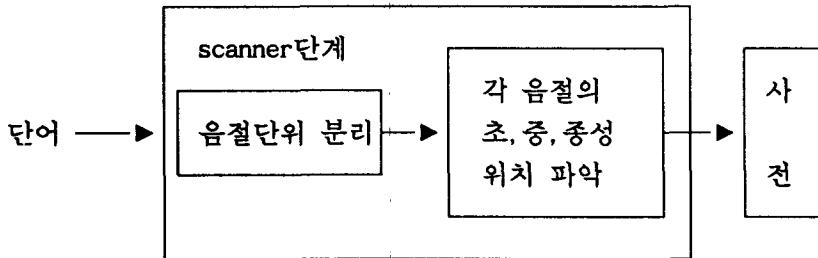
전처리 단계는 한글이 모아쓰기를 함으로 음절상에 나타나는 각 초성, 중성, 종성에 대한 위치 정보를 한글 신경망 모델에 제공한다.

컴퓨터상에 나타나는 한글은 보통 2byte완성형 및 조합형을 사용하나 본 연구에서는 2byte조합형을 사용하였다. 한글 2byte조합형은 [그림 6-2]와 같은 구조를 갖고 있다.



[그림 6-2] 2 byte 한글코드

15bit는 영어/한글을 나타내며 14 - 10 bit는 초성, 9 - 5 bit는 중성, 4 - 0 bit는 종성을 나타낸다. 2byte로 표시된 한글을 5 bit씩 bit-masking을 하여 한음절의 초성, 중성, 종성의 위치를 판별할수 있으며, 초성일 경우 1, 중성은 2, 종성은 3으로 그 위치적 정보를 나타낸다.



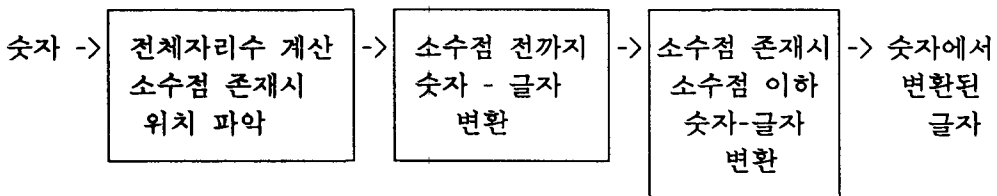
[그림 6-3] scanner 흐름도

2) 숫자 - 글자 변환 구현

한글 텍스트상에 나타나는 숫자의 경우, '1,123'을 '천 백 이십 삼'으로 읽고, 숫자 '0.12'와 같이 소수점이 포함되어 있을 때는 '영점 일 이'등으로 읽게된다. 그러므로 소수점과 콤마(,)등이 구별되어야 한다.

숫자 - 글자 변환 단계 구현시 '6.29'와 같이 특수한 경우는 제외 시켰으며 숫자 3자리수 마다 나타나는 ','가 있건 없건 간에 숫자의 자리수에 맞추어서 읽게 하였다. 즉 '1,234'와 '1234'는 똑같이 '천 이백 삼십 사'로 읽게된다. 그리고 소수점 이하의 숫자는 '영점 일 이'와 같이 숫자 하나 하나를 끊어서 읽게 하였다.

숫자 - 글자 변환 단계의 흐름도는 [그림 6-4]와 같다.



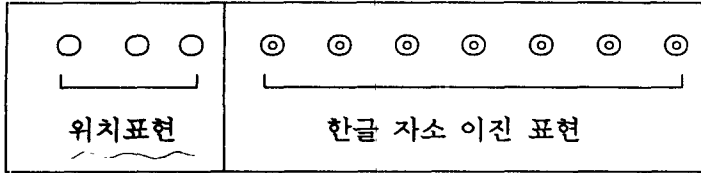
[그림 6-4] 글자 변환 단계 흐름도

3) 규칙 상충 처리 단계 구현

한 음운규칙과 상충되는 예외적인 단어들은 규칙 상충 처리 사전(dictionary)으로 구성 하였다. 이러한 조건에 해당하는 단어들의 표준형태와 실제 발음되는 내용을 함께 저장해둔다. 즉 '의견란'이라는 단어와 실제 발음 '의견난'을 같이 사전에 저장 하였다. 예외사항에 속하는 각 단어의 첫음절의 초성을 기준으로 사전적인 순서로 사전에 저장 한다.

사전 구성 흐름도는 [그림 6-5]와 같으며 사전을 [그림 6-5]와 같이 구성함으로써 단어탐색 시간을 줄일수 있으며 포인터를 사용하였을 경우 새로운 예외사항 추가시

각 자소의 코드는 이진코드 체제를 사용하였다. 즉 모두 52개이므로 7bit으로 각 자소를 표현하였다.



[그림 6-6] 입력창 구조

입력층은 5개의 입력창으로 구성되며 출력층은 모두 10개의 node로 구성된다. 출력층은 입력층과 같이, 3개의 node는 위치정보를 나타내며 나머지 7개 node는 입력층의 왼쪽으로 부터 3번째 자소에 대한 실질 음소를 출력한다.

은닉층은 훈련과정에서 음운 변동 규칙의 발견자 역할을 하며 본 연구에서는 30개의 node로 은닉층을 구성하였다.

이와같이 구성된 신경망 모델은 출력층에서 나타나는 실제 출력값과 목표값 사이의 차를 역전파하여 실제 출력값이 바라는 목표값으로 수렴하여 전체 오차합이 임계치 값보다 작게 될때까지 훈련시키는 Back propagation 학습 알고리즘[13]을 사용하였다.

6-3. 실험결과 분석

1) 실험데이터

실험 데이터는 크게 A, B, C, D 4개의 그룹으로 분류하였다. 먼저 설측음화에 해당하는 단어를 표준 국어사전[1]에서 초성이 'ㄱ'에서 'ㅎ'까지 37개 단어를 선정하였다. 이중 23개의 단어를 훈련용 데이터(A)로 사용하였고 나머지 14개의 단어(B)를 훈련된 신경망의 성능을 테스트하기 위한 분석자료로 사용하였다. 설측음화에 예외적이면서 설측음화에 상충적인 규칙을 갖는 9개의 단어를 한글 맞춤법 통일안[6] 표준 발음법에서 수집하였으며 이중 5단어(C)를 훈련용 데이터로 사용하였고 나머지 4단어(D)를 훈련된 신경망의 성능을 실험하기 위한 실험데이터로 사용하였다.

2) 실험방법

본 연구에서 제안한 한국어 글자-음소 변환 시스템을 실험하기 위해 한국어 음운 변동 규칙에 해당하는 단어와 음운변동 규칙과 상호 상충되는 예외적인 단어들을 함께 신경망에 학습 시켰을 경우의 결과와 음운변동 규칙에 해당하는 단어를 학습시키고 학습시킨 규칙적 단어와 서로 상충되는 단어들을 규칙 상충 처리 사전을 도입했을 경우를 비교하기 위해 2가지로 나누어서 실험하였다.

3) 실험 1

첫번째 실험에서는 설측음화에 해당하는 단어와 설측음화에 상충되는 예외적인 단어들을 같이 신경망에 학습시켰을 경우, 신경망의 학습효율 및 성능을 알아 보았다.

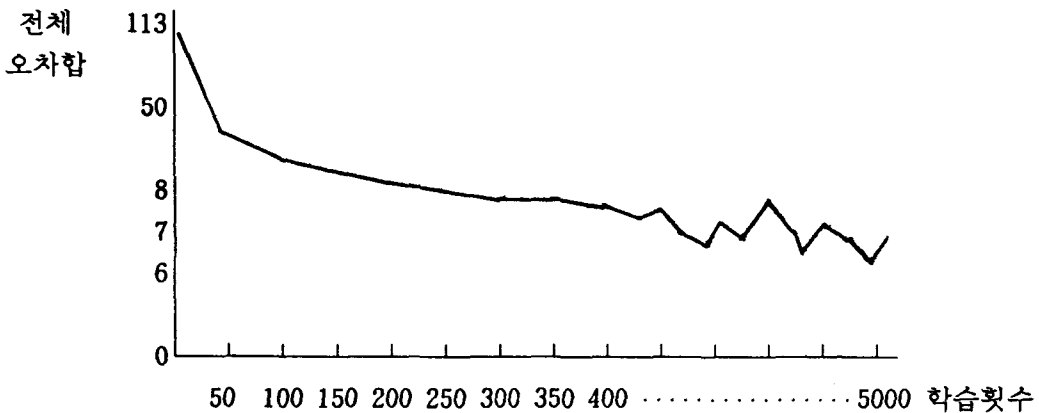
A그룹에 해당하는 23 단어와 C그룹에 해당하는 5 단어를 신경망을 학습시키는 학습 데이터로 사용하였으며 A, C그룹의 총 28 단어로 학습된 신경망의 성능, 즉 올바른

음소를 만들어 내는가를 테스트하기 위해 B그룹의 14단어와 D그룹의 4단어를 테스트 데이터로 사용하였다.

신경망의 학습 정도는 총 28단어가 충분히 학습될수 있도록 학습 횟수를 5,000번 수행 하였으며 그때까지의 전체 오차합을 학습횟수 10번마다 관찰하였다.

학습 초기에 전체 오차합 E는 112.3181이었으며 50번 학습후 신경망의 전체 오차합은 8.0012까지 감소 하였다. 그러나 학습횟수가 점점 증가 할수록 6.5068에서 4.9853, 4.9853에서 다시 6.0056으로 전체 오차합이 진동함(Oscillation)을 발견할수 있었다. 학습횟수가 5,000번 될때까지 전체오차합은 오차 한계치 2.0에 수렴하지 못하고 계속 진동하였다.

이와 같이 전체오차합이 학습횟수가 증가하더라도 수렴하지않고 계속 진동하는 이유는 바로 설측음화에 해당하는 단어와 설측음화에 상충되는 예외적인 단어를 같이 학습 시켰기 때문이다. 실험 1에서의 전체 오차합 E와 학습횟수와의 관계를 그래프로 나타내면 [그림 6-7]과 같다.



[그림 6-7] 전체 오차합과 학습횟수 비교

학습되는 과정을 살펴보면 설측음화에 해당하는 단어들은 좋은 학습율을 보인 반면 설측음화에 상충되는 예외단어들은 학습이 되지 못했다.

학습이 충분히 된 신경망에대해 B그룹 단어와 D그룹 단어로 테스트한 결과 B그룹의 데이터는 설측음화가 적용된 정확한 음소로 100% 변환되었으나 D그룹 단어들은 모두 설측음화에 예외적인 단어들 임에도 불구하고 설측음화가 적용된 음소로 출력되었다

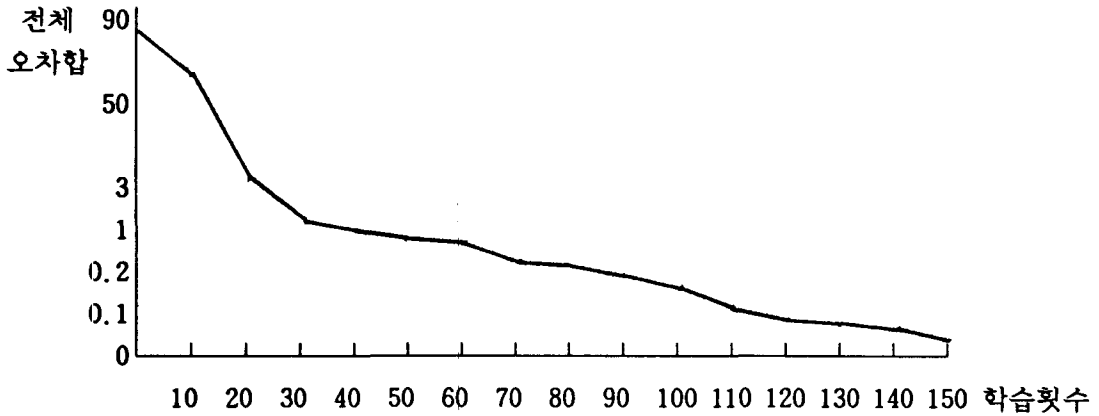
4) 실험 2

두번째 실험에서는 A그룹에 해당하는 단어 23개를 신경망의 훈련 대상으로 삼았으며, B그룹에 해당하는 단어를 성능 분석을 위해 사용하였다.

신경망의 학습 정도는 전체 오차합 E ($1/2 \sum_p \sum_i (d_{ip} - y_{ip})^2$)가 오차 허용치 0.01보다 작을때 까지 학습 시켰다.

학습이 시작될때 전체 오차합은 88.244911이었으며 152번 학습한 후 오차 한계치

0.01보다 작아짐을 알수 있었다. 학습 횟수와 전체 오차합과의 관계를 그래프로 나타내면 [그림 6-8]과 같다.



[그림 6-8] 전체 오차합과 학습횟수 비교

학습 초기에는 각 자,소의 초성, 중성, 종성의 위치를 빠르게 신경망이 학습하였다. 그러나 한글에서 초성 'ㅇ'의 음가가 없음을 학습횟수 55회 이후부터 신경망이 학습하였다.

위와같이 학습된 신경망에 대해 이미 학습된 A그룹 데이터를 입력으로 실험한 결과, 각 단어에 대해 설측음화 규칙이 적용된 100% 정확한 음소로 변환되었다. 그리고 학습되지 않은 B그룹의 자료를 입력했을 경우도 역시 설측음화가 적용된 100% 정확한 음소로 변환되었다.

7. 결 론

본 연구는 규칙기반과 신경망 모델을 결합한 한글-음소 변환 시스템 모델을 제안하고 시스템을 구현하였다. 시스템 개발을 위해 한국어 음운 변동규칙 중 설측음화를 실험 대상으로 삼았으며 50개의 node로 구성된 입력층, 30개의 node로 구성된 은닉층, 10개의 node로 구성된 출력층을 갖는 간단한 신경망 모델로 글자-음소 변환에 대해 실험 하였다.

실험2를 통해 Back-Propagation 신경망 모델을 이용하여 한국어 음운 변동규칙에 해당하는 단어를 신경망에 학습 시켰을 경우, 신경망은 적은 학습횟수로 한국어 음운 변동 규칙을 학습하였다. 그러나 실험 1을 통해 한국어 음운변동 규칙에 해당하는 단어와 그와 상충되는 단어를 같이 훈련 시켰을 경우 신경망이 제대로 학습이 되지 않았으며 학습횟수도 실험2에 비해 30배가 넘었다. 그러므로 음운변동 규칙에 상충되는 예외사항의 경우는 본 시스템에서 제안한 것과 같이 규칙 상충 처리 사전을 도입함으로써 신경망에서 잘 처리하지 못하는 부분을 규칙 기반적 접근방법으로 처리 하도록 함이 좋다.

본 연구의 한계는 한글을 음소로 변환시키는데 있어서 문장 단위로 학습시키지 못

하고 단어 단위로 학습 시켰으며 설측음화 이외의 변동 규칙과 숫자에서 글자로 변환된 데이터를 학습시키지 못했으며 변이음 처리는 하지 않았다.

앞으로 연구 과제로는 한국어 음운변동 규칙에 대해 보다 깊게 연구함으로써 신경망에 학습시킬 학습 데이터와 규칙 상층 처리 사전에 들어갈 데이터를 수집하는 것이 필요하며 이와 더불어 한글 문장을 입력으로 받을 경우, 띄어 쓰기 정보를 처리하는 node를 본 연구에서 제안한 신경망에 추가적으로 도입하여 구성함으로써 보다 좋은 성능을 갖는 한국어 글자-음소 변환 시스템을 구축할 수 있을 것이다.

* 참고 문헌 *

- [1] 민중서관, 국어사전, 1981
- [2] 박 민용, 최 항식, 뉴로 컴퓨터, 대영사, 1991
- [3] 양 재우, 김 두현 , 한글단어를 발음기호로 변환시키는 인공신경망에 관한 연구, 인공지능 소식, 11, 1988, p.32-38
- [4] 양 영중, A Study on the Development of Hangul Text-Voice Converting System, SERI, KIST, 12, 1989
- [5] 이 용주, 김 창주, A Study on the Development of Text-to-Speech Conversion Technique for Communication Processing. 한국 전자통신 연구소, 6, 1990
- [6] 이 은정, 한글 맞춤법 · 표준어 해설, 대제각, 1988
- [7] 허 용, 국어 음운론, 샘문화사, 1985
- [8] D.E. Rumelhart, G.E. Hinton, and R. J. Williams, " Learning internal Representations by Error Propagation" in D.E. Rumelhart & J.L. McClelland(EDS), Parallel Distributed Processing : Exploration in the Microstructure of Cognition. Vol1 : Foundations, MIT Press
- [9] Geoffrey E. Hinton, Connectionist Learn Procedures, Arpanet : hinton%ai. to-ronto edu@relay.cs.net, 12 1987, p.1-32
- [10] Julia B. Hirschberg, Stephen A. Riederer, James E. Rowley, and Ann K. Syrdal , VOICE RESPONSE SYSTEM : TECHNOLOGIES AND APPLICATIONS, AT&T TECHNICAL JOURNAL SEPTEMBER/OCTOBER, 1990 P.42-51
- [11] Kaszuo HAKODA, Kiyoshi KABEYA, Tatsuya HIRAHARA and Keiichi NAGAKURA, JAPANESE TEXT TO SPEECH SYNTHESIZER BASED ON RESIDUAL EXITED SPEECH SYNTHESIS, IEEE, ICASSP, 86, TOKYO, P. 2431-2434
- [12] Micheal H. O'Malley, Text-To-Speech Conversion Technology, IEEE, COMPUTER, AUGUST, 1990, P.17-23
- [13] Richard P. Lippmann, An Introduction to Computing with Neural Nets, IEEE ASSP. MAGAGINE, 1987, P.4-22
- [14] T. Sejnowski and C.R. Rosenberg, " NETtalk : A Parallel Network That Learns to Read Aloud", Johns Hopkins Univ. Technical Report JHU/EECS-86/01, 1986, P. 663-672
- [15] WALTER SCHINDE, Connectionism : Is it a paradigm shift for psychology ?, Psychonomic Society, 1987, p. 73-83