

한국어 연속음성 인식을 위한 발음열 자동 생성

Automatic Generation of Pronunciation Variants for Korean Continuous Speech Recognition

이 경 님*, 전 재 훈**, 정 민 화*

(Kyong Nim Lee*, Je Hun Jeon**, Min Hwa Chung*)

*서강대학교 컴퓨터학과 음성언어처리연구실, **삼성전자 디지털미디어총괄 컴퓨터시스템사업부
(접수일자: 2000년 12월 22일; 채택일자: 2001년 1월 29일)

음성 인식이나 음성 합성시 필요한 발음열을 수작업으로 작성할 경우 작성자의 음운변화 현상에 대한 전문적 언어지식을 비롯하여 많은 시간과 노력이 요구되며 일관성을 유지하기도 쉽지 않다. 또한 한국어의 음운 변화 현상은 단일 형태소의 내부와 복합어에서 결합된 형태소의 경계점, 여러 형태소가 결합해서 한 어절을 이룰 경우 그 어절 내부의 형태소의 경계점, 여러 어절이 한 어절을 이룰 때 구성 어절의 경계점에서 서로 다른 적용 양상을 보인다. 본 논문에서는 이러한 문제를 해결하기 위해서 형태음운론적 분석에 기반하여 문자열을 자동으로 발음열로 변환하는 발음 생성 시스템을 제안하였다. 이 시스템은 한국어에서 빈번하게 발생하는 음운변화 현상의 분석을 통해 정의된 음소 변동 규칙과 변이음 규칙을 다단계로 적용하여 가능한 모든 발음열을 생성한다. 각 음운변화 규칙을 포함하는 대표적인 언절 리스트를 이용하여 구성된 시스템의 안정성을 검증하였고, 발음사전 구성과 학습용 발음열의 유용성을 인식 실험을 통해 평가하였다. 그 결과 표제어 사이의 음운변화 현상을 반영한 발음사전의 경우 5-6% 정도 나은 단어 인식률을 얻었으며, 생성된 발음열을 학습에 사용한 경우에서도 향상된 결과를 얻을 수 있었다.

핵심용어: 연속음성인식, 한국어 음운변화 현상, 음운 모델링, 발음열 생성기, 발음사전 구축, 의사 형태소

투고분야: 음성처리 분야 (2.5, 2.7)

Many speech recognition systems have used pronunciation lexicon with possible multiple phonetic transcriptions for each word. The pronunciation lexicon is often manually created. This process requires a lot of time and efforts, and furthermore, it is very difficult to maintain consistency of lexicon. To handle these problems, we present a model based on morphophonological analysis for automatically generating Korean pronunciation variants. By analyzing phonological variations frequently found in spoken Korean, we have derived about 700 phonemic contexts that would trigger the multilevel application of the corresponding phonological process, which consists of phonemic and allophonic rules. In generating pronunciation variants, morphological analysis is preceded to handle variations of phonological words. According to the morphological category, a set of tables reflecting phonemic context is looked up to generate pronunciation variants. Our experiments show that the proposed model produces mostly correct pronunciation variants of phonological words. Then we estimated how useful the pronunciation lexicon and training phonetic transcription using this proposed systems.

Key words: Continuous speech recognition, Korean pronunciation variation, Phonetic and phonological modeling, Phonetic transcription, Pronunciation modeling, Pseudo-morpheme

Subject classification: Speech signal processing (2.5, 2.7)

I. 서 론

음성 인식이나 합성에서는 문자 형태의 발음처리가 아니라 실제 발음 형태의 발음처리가 사용되어진다. 음성인식 시스템에서는 각 단어에 대해 가능한 발음열을 모두 고려

한 것이 일반적으로 인식률이 우수하며, 음성 합성에서도 청자에게 더 친근감을 느끼도록 하기 위해 발성 상황에 따라 여러 가지 형태의 발화가 필요하다. 이러한 발음열은 수작업으로 작성할 경우 작성자가 한국어 음운 변화에 대한 전문적인 지식을 가지고 있어야 하며, 작성에 있어서도 많은 시간과 노력이 필요하다. 또한 다수의 작성자에 의해 작성되는 경우 해당 발음열에 대한 일관성 유지가 힘들다. 이러한 이유 때문에 최근 들어 발음열을 자동으로 생성하는 문제에 관해 많은 연구가 이루어지고 있다[1][12].

책임저자: 이경님 (knllee@nlpeng.sogang.ac.kr)
127-742 서울 광역시 마포구 신수동 1번지 서강대학교
컴퓨터학과 음성언어처리연구실
(전화: 02-712-8023; 팩스: 02-704-8273)

정확한 발음열을 생성하기 위해서는 해당 언어가 가지는 특징뿐만 아니라 그 언어의 음운 변화 규칙과 그 규칙의 적용 양상에 대한 정확한 분석이 필요하다.

한국어는 음절 기반의 표음 문자이기 때문에 각각의 철자는 특정 음소를 나타낸다. 그러나 실제 발음할 때에는 철자에 나타난 음소열이 그대로 발음되지 않고, 음운 변화 과정을 통해서 실제와 다른 음소의 배열을 가지는 경우가 많다. 이러한 변화는 크게 두 가지 조건에 영향을 받는다.

첫째는 음소 연결 제약 조건이다. 일부 음소가 음절의 특정 위치에 나타나지 못하거나 특정한 음소의 배열이 허용되지 않는다는 것이다. 음절이 연결되어 음소 연결 제약에 위배되는 음소의 배열이 생겨나면 이를 시정하기 위한 음운 현상이 일어난다.

둘째는 연결 내 구성 형태소의 종류와 연결 형태이다. 연결은 끊어읽기 단위로서 연결 내의 음소 배열이 같더라도 구성 형태소의 종류가 다르면 음운 변화는 다른 방향으로 일어난다. 예를 들어 '감기'라는 음소 배열에서, 그 어절이 명사로 구성된 경우 /K AA M G IY/ (감기)로 발음되고, 어간과 어미의 결합으로 구성된 경우 /K AA M KK IY/ (감끼)로 발음된다. 두번째 음절 초성 /ㄱ/이 형태소 특성에 의해 /KK/ 혹은 /G/로 달리 발음되며, 첫 음절 초성 /ㄱ/의 경우 또한 유성음 중성을 가지거나 중성이 없는 어절의 뒤에 연결로 묶여 사용될 경우 /K/ 또는 /G/로 발화되기도 한다. 다른 예로 명사와 조사가 결합한 '숨이'라는 어절과 복합어 '숨이불'이 발음될 경우, 첫번째 음절과 두번째 음절 사이의 음운변화에서 '숨이'는 /S OW M IY/ (소미)로 발화되고 '숨이불'은 /S OW M NI IY B UW L/ (숨니불)로 서로 다른 발음열을 가지게 된다. 같은 형태의 음소 배열이지만 구성 형태소가 다르므로 첫번째는 연음이 된 경우이고, 두번째는 ㄴ-첨가가 된 경우이다. 또한 '겨울 나그네' (겨울라그네)와 '너는 산을, 나는 바다를' (사늘 나는)이란 경우와 같이 하나의 연결로 묶여 발음되는 경우에는 음소 연결 제약 조건이 적용되어 /ㄹ/과 /ㄴ/이 연이어 나오면 뒤의 /ㄴ/이 /ㄹ/로 바뀌지만, 그렇지 않은 경우 아무런 변화도 없다.

마지막으로 하나의 문자에 대해 여러 발음이 존재할 수 있다. 일반적으로 이중 모음이 단모음으로 발음되는 경우를 제외하더라도 '학교'가 /H AA KQ KK JO/ (학꼬)라고 발음될 수도 있고 /H AA KK JO/ (하꼬)라고 발음될 수 있다. 이 경우 두번째 발음이 표준 발음이 되고 첫번째 발음이 비표준 발음이 된다. 하지만 실생활에서는 첫번째 발음이 종종 사용되어지고 있다.

본 논문에서는 이러한 한국어의 특징을 고려하기 위해 형태음운론적 분석을 기반으로 문자열을 자동으로 발음열로 변환하는 발음 생성 시스템을 제안한다.

II. 텍스트 전처리

일반적으로 발음열 자동 생성기의 역할은 입력된 한글 철자에 대해 발음 나는데로 표기된 음소열로 출력하는

것이다. 그러나 실제 입력되는 문장에는 특수기호 및 알파벳, 아라비아 숫자, 알파벳과 숫자가 혼용된 약어, 영문 단위명사 등이 포함되어 있다. 게다가 아라비아 숫자의 경우 문맥에 따라 숫자를 읽는 방식이 달라지기 때문에 발음열 생성기 내부에서 이 모든 현상을 수용하여 처리하는 것은 주된 목적 수행을 위한 선행 작업이 상대적으로 크게 되는 문제가 발생된다. 따라서 시스템 생성에 맞는 입력 형식으로 변환하는 작업과 형태음운론적 특성을 반영하기 위한 형태소 분석 수행을 포함하는 전처리 시스템을 구성하였다.

2.1. 특수문자 및 2byte 코드 변환

문장에서 제거 가능한 코드문자는 제거하고, 「 」 「 」 와 같은 괄호문자와 ℃, m²와 같은 기호 단위, 2바이트 로마자 숫자 II, III 등은 아라비아 숫자로, 그 외 2바이트 코드는 해당 1바이트 아스키 코드로 변환작업을 수행한다.

2.2. 형태소 태깅

문자기반의 형태소 분석기의 경우, 원형을 복원하는 루틴이 포함되어 용언의 불규칙 처리나 생략된 조사와 축약된 형태까지 복원하게 된다. 그러나 인식이나 합성에 있어서 소리값을 유지하기 위해 문자기반의 형태소 분석기를 수정하여 원형 복원하지 않도록 형태소 태깅을 수행하였다.

2.3. 기호, 단위 변환

su (단위기호), nbu (단위성 의존명사)로 태깅된 경우, 알파벳 기반 단위명사는 발화되는 한글 단위명사로 변환한다. 예를 들어 기호 'kg'은 '킬로그램'으로 변환 작업을 수행한다.

기타 발화되지 않는 구분기호로 어절 경계에 변화가 없는 경우, 기호와 태그만 제거하고 좌우의 단어를 결합하여 하나의 어절로 구성하였다. sp의 경우는 쉼표에 해당되는데, 일반적으로 쉼표 바로 뒤는 띄어쓰기가 실현되나 대용량 텍스트 처리시 띄어쓰기가 많이 생기므로 전처리시 sp 태그를 제거하고 어절 경계를 주었다. 마지막으로 sd는 이음표로, 일반적인 '·'기호의 경우는 sp와 같은 처리를 하였고, '~'의 경우는 숫자와 연결시 '에서'로 변환하고 어절 구분을 하였다.

2.4. 외국어 및 고유명사 변환

알파벳과 숫자가 섞인 고유명사 처리 문제로 대부분 외래어나 외국어, 제품명과 같은 고유명사에서 많이 발생한다. 기본적으로 알파벳과 숫자는 한 character 단위로 일대일 매칭하고, 그 외의 경우는 예외 발음처리하였다. 예외적인 경우에는 "펜티엄III (펜티엄쓰리), 3COM (쓰리콤), MP3 (엠펙쓰리), FIFA (피파), Internet (인터넷)" 등이 있다. 이 외에도 "3.1 (삼일), 12.12 (십이십이)" 등과 같은 숫자 고유명사의 경우 예외 처리를 수행하여 변환하였다.

2.5. 숫자처리 및 변환

아라비아 숫자의 경우는 뒤에 오는 단위명사에 따라 '일이' 식이나 '한두' 방식으로 발화하게 되므로 단위명사 별로 분류하여 처리하였다. '한두' 방식으로 읽게 되는 단위명사 목록으로는 대표적으로 "시간, 건, 골, 바퀴, 살, 번째, 포기, ..." 등이 있으며 100단위 이하의 경우에만 적용하였다. 분류된 단위명사의 경우 화자의 습관에 따라 '일이' 식으로 혼용되어 읽는 경우도 발생하여 음성과 일치하지 않는 오류가 전처리기에서 발생할 수도 있다. 또한 그 외에 소수점 읽기 및 숫자 읽기 단위 등을 처리하도록 하였다.

2.6. 의사형태소 태깅

의사형태소 태깅은 접속정보 규칙[3]에 따라 형태소 태깅 결과로부터 자동 재태깅을 수행하였다. 그 외 발음열 생성 측면에서 추가된 사항을 정리하면 다음과 같다.

파생어 특히 '~적, ~성, ~권' 등과 같은 접미사가 붙은 한자어의 경우 경음화 처리 문제가 발생한다. 이는 국어학적으로도 도출 가능한 규칙이 정해져 있지 않기 때문에 대부분 예외처리를 할 수 밖에 없다. 예외사전을 이용한 예외처리를 수행하기 위해 해당 접미사의 경우 일반명사로 결합하는 규칙이 추가되었다. 또한 복합어 관계에서 발생하는 L-첨가 규칙을 적용하기 위해 복합명사 사전에 복합어 리스트를 기록하고 검색을 통해 복합명사 태그를 주도록 구성하였다. 현재 등록된 복합명사 리스트는 총 67개이다.

III. 한국어의 형태 음운학적 특징

한국어 단어의 철자는 음운 변화 과정을 통해 발음열로 바뀌게 된다. 음운 변화 과정은 음운 변화가 일어나는 위치에 따라 음소 변동 규칙과 변이음 규칙을 사용해서 설명할 수 있다. 해당 음소 문맥에 의해 하나의 음소가 다른 음소로 바뀌거나, 탈락하거나, 첨가되는 양상을 규칙화한 것을 음소 변동 규칙이라 하고, 하나의 음소가 발화 상에서 여러 변이음으로 실현되는 양상을 규칙화한 것을 변이음 규칙이라 한다[6]. 음소 변동 규칙은 규칙의 적용 양상에 따라 다시 필수 음소 변동 규칙과 수의적 음소 변동 규칙으로 나누어진다. 그림 1은 '학교'라는 단어가 음운 변화 과정을 통해 실제 발음열로 바뀌는 과정을 나타낸 예이다. '학교'라는 단어는 우선 필수 음소 변동 규칙을 받아 /학교/라는 발음을 가지고, 다시 수의적 음소 변동 규칙을 적용 받아 /하교/라는 발음을 가지며, 최종적으로 변이음 규칙을 적용 받아 /H AA KQ KK JO/ 또는 /H AA KK JO/라는 발음열을 가지게 된다. 이러한 음운 변화 과정은 언절의 구성 형태소의 구성에 따라 적용 양상이 달라진다. 형태음운론적 분석이란 발음열 생성에서 음운 변화 과정이 언절을 이루는 형태소의 구성에 따라 다르게 적용되는 양상을 분석하는 것을 의미한다[1]. 하나 이상의 형태소가 모여 어절을 형성하고 이 어절들이 모여 언절을 형성하며 이 언절들이 실제 발화로 이어진

다. 따라서 같은 형태의 음소배열을 가지더라도 실제 발화 시에는 언절을 이루는 각 형태소의 구성에 따라 여러 가지 다른 모양의 발음이 나타나게 되는 것이다. 본 논문에서는 보다 정확한 발음열 생성을 위해 음소 문맥에 따라 필수 음소 변동 규칙, 수의적 음소 변동 규칙, 변이음 규칙으로 정의하고, 언절의 구성 형태소의 변화에 따른 음운 변화 규칙이 적용되는 양상을 규명하였다

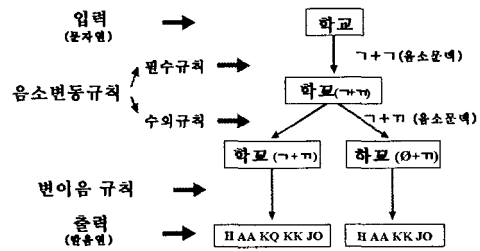


그림 1. 음운변화 과정 예제 Fig. 1. An example of phonemic variation process.

3.1. 음소 변동 규칙

음소 변동 규칙을 적용해서 변화되는 음소와 변화에 영향을 주는 음소의 배열을 음소 문맥이라 한다. 일반적으로 음소 문맥은 연속으로 나오는 두 자음의 순서쌍으로 나타난다. 즉 음소 변동이 일어나는 음절 경계의 앞 음절 끝 자음과 뒤 음절 첫 자음의 쌍으로 이루어진다. 그러나 구개음화나 L-첨가의 경우는 뒤 음절의 중성도 음소 문맥으로 고려되어야 하고, 모음화 규칙들과 모음 탈락의 경우는 뒤 음절 중성과 앞 음절의 중성까지도 음소 문맥에 포함되어야 한다.

다음의 표 1은 음소 문맥을 바탕으로 장애음의 비음화 규칙의 세부 규칙 중 일부를 표현한 예를 나타낸 것이다. 음소 문맥항의 L3는 음소 변동이 일어나는 음절 경계의 앞 음절 종성을 나타내고, R1은 뒤 음절의 초성을 나타낸다. 변환 코드는 해당 음소 문맥에 대한 음소의 변동 결과를 나타낸다. 적용범위는 세부규칙의 적용 양상을 나타낸다. 여기서 (여)는 여러 어절이 한 언절을 이룰 때 구성 어절간의 경계점에서, (형)은 여러 형태소가 결합하여 한 어절을 이룰 경우 구성 형태소의 경계점에서, (복)은 복합어에서 결합된 형태소간의 경계점에서, (내)는 단일 형태소의 내부에서 규칙 적용 여부를 나타낸다. 이때 '1'은 적용됨을, '0'은 적용되지 않음을 의미한다. (수)는 규칙의 적용이 수의적으로 이루어짐을 나타낸다. 같은 음소 문맥에 둘 이상의 음소 규칙이 대응되는 경우에는 틀린 발음열을 생성하지 않도록 규칙의 우선 순위를 정의하여야 한다. (다)는 하나의 음소 문맥에서 두 개 이상의 규칙이 동시에 적용될 수 있는지 아니면 하나만이 적용되어야 하는지를 나타내는 부분으로, 두 개 이상의 규칙이 적용될 경우 정의된 우선 순위를 이용하여 규칙을 적용할 수 있도록 한다. '1'인 경우 여러 규칙이 동시에 적용될 수 있음을 나타내고, '0'인 경우는 하나만 적용되어야 함을 의미한다

표 1. 장애음의 비음화 규칙의 세부규칙 표현 일부
Table 1. Representation of Nasalization of obstruent rule.

음소 문맥		변환 코드		규칙 번호	세부 규칙 번호	적용범위 어/형/복/내/수/다
L3	R1	L3	R1			
ㄱ	ㄴ	→	ㅇ	6	1	111100
ㄱ	ㄹ	→	ㅇ		2	011100
ㄱ	ㄷ	→	ㅇ		3	111100
ㄱ	ㄸ	→	ㅇ		4	111100
ㄱ	ㄹ	→	ㅇ		5	111100

한국어의 음소 변동 규칙 중 발음열 생성에 필요한 규칙은 표 2와 같이 총 17개의 규칙으로 정의하였다[1][6][9]. 기존의 논문[8]에서는 20개의 음소 변동 규칙을 정의했었으나, 이 중 비음동화 규칙은 세부규칙의 구현 과정에서 변자음화에 포함되었고, 음운 전위는 모든 음소 문맥에 적용 가능하지만 그 결과는 음바르지 않은 경우가 대부분이므로 구현을 할 필요성이 없는 규칙으로 판단하여 제외하였다. 또한 어간 어미의 반모음화와 모음 탈락 규칙은 이형태를 인정하는 형태소 태깅을 수행하므로 규칙에서 제외하였고, 전설모음화의 경우 대부분 이미 음운현상이 텍스트에 반영되고 있으므로 규칙에서 제외하였다. 일부 음운 전위가 적용되는 어절은 그 수가 적고 한정적이므로 예외 사전에 정의하여 수용할 수 있도록 문제를 해결하였다. 따라서 본 논문에서는 표 2와 같이 17개의 규칙만을 정의하였다. 17개의 규칙 가운데 입력 음소열에 대해 그 규칙에 해당 음소 문맥이 있는 경우 반드시 적용되어야 하는 필수 음소 변동 규칙은 규칙 1항에서 규칙 11항, 규칙 16항, 그리고 규칙 17항으로 총 13개이다.

표 2. 음소 변동 규칙의 분류 및 세부규칙 수
Table 2. List of phonemic rules and the number of phonemic contexts.

규칙 번호	규칙의 종류	세부 규칙수
1	음절말 중화	117
2	자음군 단순화	256
3	격음화(기식음화)	21
4	연음규칙	42
5	유음화	10
6	장애음의 비음화	34
7	유음의 비음화	19
8	구개음화	3
9	경음화	136
10	중성 ㅇ-탈락	1
11	ㄴ-첨가	30
12	동일 조음위치 자음탈락	7
13	중복 자음화	6
14	변자음화	17
15	초성 ㅇ-탈락	5
16	자음 첫소리 '의' 모음화	18
17	용언의 활용형 '저, 쟈, 처'의 모음화	3

이 중에서 규칙 11항은 복합어의 경계점과 어절의 경계점에서만 적용되며, 규칙 16항과 규칙 17항은 용언의 어간과 어미의 경계점에서만 적용된다. 그 외의 규칙은 모

든 경계점에서 대부분 적용된다. 이와는 달리 규칙 11항에서 규칙 15항까지의 4개의 규칙은 수의적으로 적용되는 규칙이다. 수의적 규칙을 적용하여 생성한 발음열은 표준 발음으로 인정되지 않는다. 그러나 이 경우의 비표준 발음이 실제 대화에서 빈번히 발생하므로 음성 인식 시스템의 인식률을 높이기 위해 표준 발음과 함께 발음사전에 정의해줄 필요가 있다.

3.2. 변이음 규칙

하나의 음소는 음성 환경, 말의 속도와 스타일에 따라서 여러 가지 음가를 가지게 되는데, 이러한 소리들을 변이음이라 하며, 하나의 음소가 여러 변이음으로 변화되는 현상을 규칙화한 것을 변이음 규칙이라 한다. 한국어에서 존재하는 변이음 규칙의 종류는 반모음의 무성화 규칙, 모음의 무성화 규칙, 구개음화 규칙, 목젓 소리되기 규칙, 유성음화 규칙, 자음 약화 규칙, 원순음화 규칙, 비음화 규칙, 무파화 규칙, 탄설음화 규칙 등이 있다[2][6][7][8].

표 3. 변이음 정의
Table 3. List of PLUs.

번호	자모	표기	번호	자모	표기	번호	자모	표기
1	ㅂ	P	19*	ㅈ	ZZ	37	ㅣ	IY
2	ㅃ	B	20	ㅊ	CH	38	ㅑ	EH
3	ㅍ	PQ	21	ㅅ	S	39	ㅓ	EY
4	ㅍ	PP	22	ㅆ	SI	40	ㅕ	OI
5	ㅍ	PH	23*	ㅈ	SS	41	ㅗ	UI
6	ㄷ	T	24*	ㅉ	SY	42	ㅜ	JA
7	ㄷ	D	25	ㅎ	H	43	ㅟ	JX
8*	ㄷ	TQ	26*	ㅎ	HO	44	ㅠ	JO
9*	ㄷ	TT	27	ㅁ	M	45	ㅢ	JU
10*	ㅌ	TH	28	ㄴ	N	46	ㅤ	JH
11	ㄱ	K	29*	ㄴ	NI	47	ㅥ	JE
12	ㄱ	G	30	ㄹ	L	48	ㅦ	WA
13*	ㄱ	KQ	31	ㄹ	R	49	ㅧ	WX
14*	ㄱ	KK	32	ㅏ	AA	50	ㅨ	WH
15*	ㅋ	KH	33	ㅑ	AX	51	ㅩ	WE
16	ㅇ	NX	34	ㅓ	OW	52	ㅪ	WI
17	ㅈ	Z	35	ㅕ	UW			
18	ㅉ	ZO	36*	ㅗ	WW			

그러나 이렇게 많은 규칙과 그에 따른 변이음을 모두 구분하여 인식하는 것은 인식 시스템을 구성하는데 있어서 PLU (Phone-Like-Unit : 유사음소)의 수만을 늘리게 되어 학습 시에 데이터 부족을 초래할 수 있고, 성능 향상에 오히려 나쁜 영향을 줄 수 있기 때문에 본 생성기에서는 유성음화 규칙, 무파화 규칙, 구개음화 규칙, 탄설음화 규칙만을 사용하였다. 유성음화는 이완 장애음 /ㅂ, ㄷ, ㄱ, ㅈ, 마찰음 /ㅈ/이 유성음과 유성음 사이에서 유성음으로 변화되는 규칙이고, 무파화는 파열음 /ㅂ, ㄷ, ㄱ/이 종성에 나타날 때 파열 단계가 생략되어 일어나는 현상이며, 탄설음화는 유음 /ㄹ/의 환경에 따른 변화 현상, 구개음화는 마찰음 /ㅅ, ㅆ/이나 비음 /ㄴ/이 모음 /ㅣ/나 혹은 반모음 /j/ 앞에서 각각 치경구개 쪽으로 이동되는 현상이다. 우리말의 발음 위치와 조음 방법을 기준으로 하여, 위에서

언급된 변이음 규칙을 적용하여 만들어진 변이음 구성은 표 3과 같이 모두 52개로 정의하였다[6][7][8]. 변이음 표기는 ARPABET의 표기법을 따랐으며 ARPABET에 정의되지 않은 음소는 새로이 정의하였으며 번호에 '**'로 표시하였다.

IV. 음운 변화 규칙의 단계적 적용

기존의 발음열 생성 방법[2]은 모든 규칙을 일괄적으로 적용하는 방법을 취하였는데, 이러한 방법은 한국어의 특징을 고려할 때 적당하지 않다. 한국어는 하나의 문자에 대해 여러 가지 발음이 가능한 경우가 많이 존재하므로 모든 발음열들을 체계적으로 생성할 수 있어야 한다. 본 논문에서는 형태소 해석 및 분석에서 널리 쓰이는 two-level 모델을 발음열 자동 생성에 적용하였다[2][5][9][10]. 여기서 two-level이란 기저형과 표면형을 의미하며, 표 4에서 보듯이 한글 자모 철자를 기저형이라 하고, 실제 철자가 주변환경의 영향으로 변화되어 나타나는 음가를 표면형이라고 정의한다. 기저형에서 표면형으로 변화하는 과정에서 음운 변화 규칙들을 단계적으로 적용하여 가능한 모든 발음열들을 순차적으로 생성하도록 하였다.

표 4. Two-level 모델 예제

Table 4. Two-level model for generating pronunciation from orthographic transcriptions.

기저형	표면형
한국	HAANXGUWKQ (한국)
싫은	SIYPHWWN (싫은)

발음열 자동 생성 알고리즘은 앞 절의 그림 1과 같이 3 단계로 진행된다[2][9] : (1)표준 발음 생성을 위한 형태소 범주에 따라 필수적 음소 변동 규칙 (표 2의 규칙 1~11, 16, 17)적용한다. 이 단계는 형태소의 범주에 따라 어간, 어미, 복합어, 조사, 명사·부사·관형사 (default), 다섯 개로 분리하여 수행한다. 각 형태소 범주와 적용되는 위치에 따라 5개의 필수 음소 변동 규칙들이 선택적으로 적용된다. 음소 변동 규칙으로 처리할 수 없는 발음열은 예외 발음사전을 두어 올바른 발음열이 생성되도록 하였다. (2)비표준 발음 생성을 위한 수의적 음소 변동 규칙 (표 2의 규칙 12~15)을 적용한다. 수의적 규칙이 적용된 발음열은 비표준 발음열이므로 필수 규칙이 적용된 발음열은 그대로 두고, 새로운 발음열을 생성하게 된다. 즉 필수 규칙이 적용된 발음열에 대해 하나 복사본을 만든 다음 그 복사본에 대해 수의적 음소 변동 규칙을 적용한다. (3)변이음 규칙의 적용한다. 변이음 규칙 적용 단계를 거치면 정의된 PLU로 구성된 발음열이 생성되어 진다

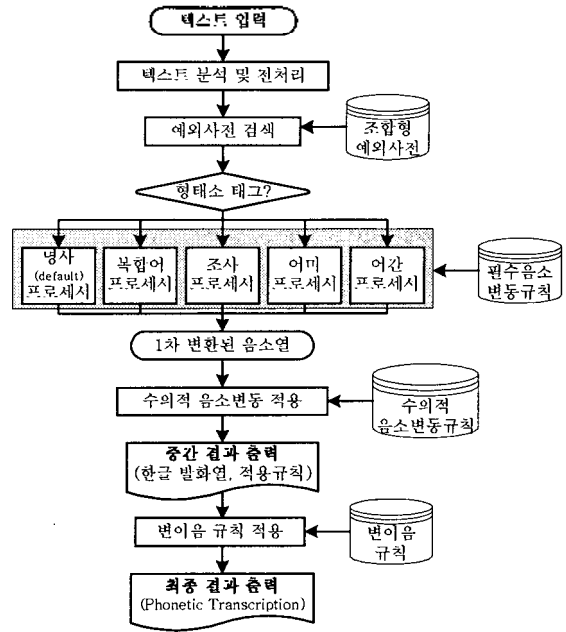


그림 2. 발음열 자동 생성 알고리즘
Fig. 2. Pronunciation generation algorithm.

4.1. 1단계 : 표준 발음열 생성을 위한 필수 음소 변동 규칙 적용

필수 음소 변동 규칙은 음소 문맥과 형태소 범주에 따라 적용되는 양상이 달라진다. 음소 문맥은 한국어의 음소 연결 제약에 의해 결정되고, 필수 음소 변동 규칙이 적용되는 형태소 범주는 형태소의 구성에 따라 네 가지로 나누어진다. 첫째는 단일 형태소 내부, 둘째는 두 개 이상의 형태소가 모여 복합어를 이룰 때 각 형태소간의 경계점, 셋째는 두 개 이상의 형태소가 모여 어절을 이룰 때 각 형태소간의 경계점, 넷째는 여러 어절이 모여 언절을 이룰 때로 구분되어진다. 음소 문맥을 형태소의 범주에 따라 구분하여 필수 음소 변동 규칙 오토메타로 정의하였다.

오토메타를 구성할 때, 일부 규칙은 반복적으로 적용되기도 한다. 자음군 단순화 후 연음이 되는 경우와 L-첨가 후 유음화되는 경우이다. 이 경우 반복적으로 오토메타를 참조하여 규칙 적용이 이루어지는데, 오토메타의 참조 회수가 일정하지 않다. 따라서 본 논문에서는 그림 3과 같이 각 오토메타 상에서 반복 적용되는 규칙을 하나로 통합하여 한 번의 오토메타 참조로 결과를 도출하도록 하였다.

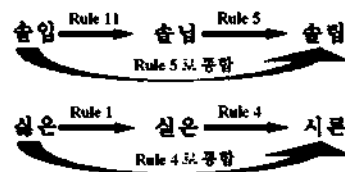


그림 3. 규칙 통합 과정 예제
Fig. 3. An example of processing of combination rules.

다음 표 5는 4개의 일부 오토메타 예제로서, 각 오토메타에서 행은 음운 변화가 일어나는 경계점의 앞 음절 종성이고 열은 뒤 음절 초성을 의미한다. 각 음소 문맥은 상태 표시와 음소 변환 코드가 있다. 상태 표시는 해당 오토메타가 음소 문맥이 3개 이상이 되도록 해주는 기능과, 규칙의 이중 적용 가능성을 나타낸다. 대표적으로 상태 m은 규칙이 이중 적용될 수 있음을 나타낸다. 예로, '감기'와 '학교'란 단어의 필수 음소 변동 규칙의 적용 과정을 보면, '감기'가 명사일 경우 표 5의 (a)를 참조하여 /감기/로 변화가 없고, 어간+어미인 경우는 (c)를 참조하여 /감끼/로 변한다. '학교'의 경우는 (a)를 참조하여 /학꼬/로 변화된다.

표 5. 필수 음소 변동 규칙 오토메타 예제
Table 5. An example of phonemic rule automata.

	ㄱ	ㄴ	ㄷ	ㄹ
ㄱ	s	n	s	s
	9.1		9.2	7.1
	ㄱㄱ		ㄱㄷ	ㅇㄴ
ㅁ	n	n	n	s
				7.8
				ㅁㄴ

(a) 형태소 내부

	ㄱ	ㄴ	ㄷ	ㄹ
ㄱ	s	n	s	s
	9.1		9.2	7.1
	ㄱㄱ		ㄱㄷ	ㅇㄴ
ㅁ	n	n	n	s
				7.8
				ㅁㄴ

(b) 복합어 경계

	ㄱ	ㄴ	ㄷ	ㄹ
ㄱ	s	n	s	s
	9.1		9.2	7.1
	ㄱㄱ		ㄱㄷ	ㅇㄴ
ㅁ	s	n	e	s
	9.121		9.122	7.8
	ㅁㄱ		ㅁㄷ	ㅁㄴ

(c) 형태소간 경계

	ㄱ	ㄴ	ㄷ	ㄹ
ㄱ	m	n	s	s
	9.1		9.2	7.1
	ㄱㄱ		ㄱㄷ	ㅇㄴ
ㅁ	n	n	n	m
				7.8
				ㅁㄴ

(d) 어절간의 경계

4.2. 2단계: 수의적 음소 변동 규칙을 적용한 다중 발음열 생성

인식이나 합성에서 사용되는 발음열은 표준 발음열 이외에도 화자의 발성 속도, 태도에 따라 발음이 달라질 수 있다. 이 단계에서는 앞 단계에서 생성된 표준 발음열의 복사본을 이용해서 또 다른 발음열을 생성한다. 정의된 수의적 규칙은 이러한 조건을 만족하는 규칙을 정의한 것으로 모든 형태소에 관계없이 적용되어진다. 표 6은 수의적 규칙을 오토메타로 나타낸 것으로, 상태 표시 b는 수의적 음소 변동 플레그로 음소 문맥을 각 항에 적용하도록 하는 것이다. 앞 단계에 생성된 단어에서 명사 '감기'는 수의적 음소 변동 규칙을 적용 받지 않으므로 변화가 없고, 어간+어미의 '감기'는 1차 변화 /감끼/와 2차 변화 /강끼/로, 명사 '학교'는 1차 변화 /학꼬/와 2차 변화 /하꼬/로 생성된다. 여기서 2차 변화 발음열 /강끼/와 /하꼬/는 발화가 빠른 경우 실현될 가능성이 큰 발음이다.

표 6. 수의적 음소 변동 규칙 일부 예제
Table 6. An example of optional phonemic rule automata.

	ㄱ	ㄴ	ㄷ	ㄹ	ㅁ
ㄱ	n	n	n	b	n
		13.1	∅	ㄴ	
ㅁ	b	b	n	n	n
	15.15	ㅇㄱ	15.16	ㅇㄱ	

4.3. 3단계: 변이음 규칙의 적용

음성 인식이나 합성에서는 음소가 아닌 변이음 단계의 발음열이 사용되어지는데, 이전 단계까지 생성된 발음열은 음소 단위의 변화이다. 현 단계는 음소 단계 표현에서 변이음 단계로 바꾸어 준다. 실제 변이음은 발음되는 음소를 기준으로 전후의 음소의 배치에 따라 변이음 특징이 결정되어진다. 생성기는 이러한 특징을 만족하도록 정의된 변이음 중 최적의 변이음을 생성한다. 명사 '감기'는 /K AA M G IY/로 어간+어미의 '감기'는 /K AA M NX KK IY/로 명사 '학교'는 /H AA [KQ] KK JO/로 실현된다. 만약 이 단어들 앞에 유성음 음소가 하나의 연결로 묶이면 /K/와 /H/가 /G/와 /HO/로 유성음화될 수 있다.

V. 실험 및 결과분석

5.1. 발음열 생성기 성능 평가

성능 평가를 위해 본 논문에서는 전형적인 음운 변화 현상을 포함하는 100개의 연결을 사용하여 기존의 시스템과 형태소 분석 결과의 이용 여부에 따른 결과를 비교하였다. 각각의 음운변동 규칙을 포함하도록 구성된 연결에 대한 발음열 생성 실험에 대한 결과는 표 7과 같다.

표 7. 발음열 생성 실험 결과
Table 7. Results of pronunciation generation.

실험 환경	입력어절	생성	미생성
기존 생성기[9] (형태소 분석 결과 고려)	162	149	13
형태소 분석 결과 미고려	108	99	9
형태소 분석 결과 고려	162	156	6
의사 형태소 분석 결과 고려	153	150	1

대부분의 연결은 한 어절로 구성되어 있으며, 어절은 대부분 2개 이상의 형태소로 구성되어 있다. 적용된 규칙의 양상을 살펴보면, 필수 음소 변동 규칙에서는 형태소 경계에서 가장 많은 규칙이 사용되었으며, 수의적 음소 변동 규칙에서는 형태소 내부에서 가장 많은 규칙이 사용되었다. 사용된 규칙 중에서는 연음 규칙과 경음화 규칙이 가장 많이 사용되었다. 과다 생성된 발음열은 두 개 이상의 수의적 규칙이 적용된 발음열이었다.

의사형태소 분석 결과를 이용한 발음열 생성 실험의 경우, 총 형태소 수는 153개, 생성된 발음열 수는 210개였다. 여기서 적용된 필수 음소 변동 규칙 수는 87개, 수의적 규칙 수는 34개, 그리고 예외 발음사전을 참조하여 생성한 발음열은 5개로 이는 규칙만으로는 생성할 수 없는 발음열이었다. 제안한 모델은 형태소 분석을 도입함으로써 보다 정확하고 완벽한 발음열을 생성해 냈다.

5.2. 예외사전의 구성

기존 예외 발음사전에 기록된 단어에 대해 본 생성기가 예외 발음을 얼마나 생성하느냐 하는 실험이다. 실험에 사용된 예외 발음사전은 한국전자통신연구원에서 제공한 예외 발음사전 중 장음을 제외한 ‘ㄱ’과 ‘ㄴ’으로 시작하는 항목의 2881어절이다. 1개의 표준 발음을 가지는 항목은 2669개였고, 2개의 표준 발음을 가지는 항목은 213개였다.

표 8. 예외 발음 생성 실험
Table 8. Exceptional pronunciation generation.

		생성된 발음열 위치				총합
		1번째	2번째	3번째	4번째	
1개 표준 발음		790	159	24	19	992
2개 표준 발음	1 발음	174	3	1	1	179
	2 발음	0	0	0	0	0

전체 생성 결과는 표 8과 같으며, 본 생성기는 기존의 예외 발음에 대해 38%정도의 생성률을 보였다. 예외 발음들 중 본 생성기가 생성한 경우는 복합어의 경계에 대한 ㄴ-첨가 규칙이 가장 많은 비율을 차지했다. 본 생성기에서 두번째 이상의 발음에서 생성한 발음열은 동일 조음위치 자음 탈락 규칙과 초성 ㅎ-탈락으로 생성된 발음이었다. 생성하지 못한 발음은 첫번째가 복합어 경계에서 의미적으로 사이시옷이 첨가되어야 하지만 앞 형태소의 종성이 존재하므로 사이시옷이 첨가되지 못하고 흔적만 남아 경음화가 되는 경우가 가장 많았고, 이중모음이 단모음으로 바뀌는 현상이 그 다음으로 많았고, 유음화 규칙에 역행하는 경우 순으로 나타났다. 첫번째 경우는 의미 분석까지 수행해야 생성할 수 있는 경우이고, 세번째 경우는 발화상의 특징으로 규칙으로는 생성하지 못하는 것이었다. 두번째 경우는 이중모음이 단순히 단모음으로 발음되는 경우로, 이중모음의 종류에 따라 단순히 단모음으로 변하는 경우도 있었으며, 이중모음과 단모음 둘로 발음되는 경우도 있고, 음절의 위치와 포함된 형태소의 종류에 따라 서로 다른 단모음으로 변하는 경우도 있었다. 이중 단순히 단모음으로 변하는 경우는 규칙화 할 수 있는 경우로 보였다. 기존 예외사전 크기는 약 25K였고,

본 시스템에서 규칙으로 생성 가능한 단어를 제거한 현재 예외사전 크기는 약 10K이다.

5.3. 한국어 연속음성 인식 시스템

연속 HMM을 기반으로 한 화자 독립 시스템으로 Entropic사의 HTK (Hidden Markov Toolkit)[13]를 사용하여 전반적인 인식 실험을 수행하였다. 본 실험에서는 16KHz로 샘플링 된 음성신호에 대해 25msec의 해밍 윈도우를 사용하여 프레임 분석을 한 후, 프레임 단위로 13차 mel-frequency cepstral coefficients, 13차 delta-cepstral coefficients, 13차 delta-delta cepstral coefficients로 구성된 특징벡터를 학습 및 인식을 위한 벡터로 사용하였다. HMM 구성은 5-state left-to-right 모델을 사용하였으며, Mixture는 기본적으로 1개를 사용하였다.

단어 인식 과정에서는 발음사전과 서브워드 모델을 사용하였다. 서브워드 모델은 각 유사 음소에 대해 학습된 HMM이며, 이 HMM은 유사 음소와 같은 모노폰의 형태에서 확장된 음소 문맥 정보를 포함하는 트라이폰으로 학습하였다. 어휘 모델의 결과인 발음사전은 인식 대상 어휘에 대한 음소 단위의 발음정보를 제공해 준다. 인식기는 입력 발화에 대해 발음사전을 참조하여 확률적으로 높은 HMM 모델 순서와 매칭이 되는 단어를 찾아 발음열을 텍스트로 복원한다. 단어인식이 이루어지면 언어모델을 이용하여 문장 인식을 수행한다. 여기서는 back-off bigram 언어모델을 사용하여 실험을 수행하였다. 음향모델로는 신문, 서적 등으로부터 추출한 낭독체 17200문장을 20대 남성화자 172명이 발성한 음성을 사용하였고, 학습용 발음열로는 청각 전사 (auditory transcription)결과를 사용하였다. 이때 변이음 규칙은 사전 크기를 고려하여, 무파화 규칙과 구개음화 규칙만을 적용하고, 유음화와 규칙은 제외하였다.

실제 인식에 사용된 테스트 문장은 학습에 사용된 문장 중 3100문장이며, 사전 표제어 단위 및 언어모델에 사용된 디코딩 단위는 의사 형태소[3]이다. 이때, 사전의 표제어 (단어 리스트)수는 약 10k이고, 한 문장당 평균 디코딩 수는 19.36개, 어절수는 9.95개이며, 테스트 문장의 perplexity는 58.72, entropy는 5.88bits였다.

5.4. 발음열 생성기를 이용한 발음사전

연속음성 인식을 수행하기 위해서는 발음사전과 언어 모델이 필요하다. 이 둘 사이에는 디코딩 단위가 일치하여야 하므로 발음사전 구축시 디코딩 단위로 표제어 단위를 선정하며 표제어 사이의 음운변화 현상을 반영한 발음사전을 구축하여야 한다. 표제어간 음운변화 현상을 반영하기 위해 본 발음열 생성기를 이용하여 발음사전을 구축하였다. 사전 구성 형식에 따라 발음열 생성기의 출력인 발음열 리스트와 사전 표제어를 정렬하여 사전을 구성하였다. 인식실험에 사용한 발음사전은 다음과 같은 특성을 지닌 사전형태로 구축하여 성능평가를 수행하였다[4].

표 9. 구축된 발음사전 특성

Table 9. Characteristics of pronunciation dictionary.

구분	사전 특성
TagVarBest	입적 언절내의 표제어간 음운변화 현상을 태그 정보를 이용하여 얻은 최적의 발음열로부터 구축된 발음사전
TagVarBest	입적 언절내의 표제어간 음운변화 현상을 태그 정보를 이용하여 얻은 다중 발음열로부터 구축된 발음사전
Standard	전문가에 의한 수작업으로 구성된 1-Best 표준 발음사전
WordBest	입적 언절내의 표제어간 음운변화를 고려하지 않고 독립된 단어 입력으로 생성된 발음사전

형태소 태그정보를 이용하여 표제어간 음운변화 현상을 반영한 최적의 발음사전과 다중 발음사전, 언어학적 기준에 의한 수작업으로 생성한 표준 발음사전, 그리고 표제어간 음운변화 현상을 고려하지 않고 생성한 발음사전을 이용하여 각각 표 10과 같은 인식 실험결과를 얻었다.

표 10. 발음사전별 인식을 및 사전크기

Table 10. Comparison of recognition rate and dictionary size.

분석결과 사전	Word Corr(%)	Word Acc(%)	Dictionary(rate)
TagVarBest	48.99	45.82	1.11
TagVarMulti	50.19	46.64	1.52
Standard	45.90	43.10	1.0
WordBest	43.21	40.99	1.0

실험결과 표제어간 음운변화 현상을 반영하지 않은 경우 단어 인식이 43.21%인 반면, 표제어간 음운변화 현상을 반영한 1-Best 사전의 경우 48.99%, Multi 사전의 경우 50.19%로 인식이 5~6% 정도 향상되었음을 볼 수 있었고, 수작업에 의한 표준 발음사전의 단어 인식이 45.90% 보다도 약 3~4% 좋은 성능을 보였다.

5.5. 학습용 발음열 생성 실험

학습 발음열이 실제 데이터베이스의 발음열과 일치하지 않게 되면, 학습과정에서 부적절한 음향으로 음소 단위들이 오염되어, 결국 이는 전체적인 성능저하를 초래하게 된다. 따라서 전체적인 음성 인식을 향상시키기 위해서는 정확한 음운변화 현상의 분석을 통한 자동 발음열 생성기가 필요하게 된다. 이 논문에서 구성된 발음열 생성기의 효용성을 평가하기 위하여 다음과 같은 실험을 통하여 비교 실험을 수행하였다.

5.3에서 학습에 사용된 사람이 직접 듣고 작성한 전사문장 (Auditory Transcription)과 발음열 생성기를 통해 작성된 발음열 문장 (Text-To-Phoneme Transcription)을 각

각 학습에 사용하여 인식 실험을 수행하였다. 실험의 객관성을 위해 발음사전은 5.4절에서 사용된 Standard 발음사전을 사용하였다.

표 11. 학습용 발음열에 따른 인식을 비교

Table 11. Comparison of recognition rate on training transcription.

분석결과 학습발음열	Word Corr(%)	Word Acc(%)	Triphone 수(uniq)
Auditory Transcription	45.90	43.10	18,967
TTP Transcription	67.53	61.67	11,359

표 11에서 보는 바와 같이 생성기로 작성된 발음열을 사용한 경우 단어 인식이 약 20% 정도 향상된 것을 볼 수 있다. 일반적으로 들리는 데로 받아 적기 작성을 하는 경우, 이미 사람이 인지적으로 인식을 통해 원본 텍스트로 변환하여 받아들이기 때문에 정확한 발음변이 현상이 반영된 전사문장을 작성하기 어려울 뿐만 아니라 일관성 유지가 어려우며, 앞뒤 음소의 영향에 의해 귀로는 한 음소의 발음을 정확하게 구별하여 듣기 어려운 문제가 발생한다. 또한 단어 경계를 찾아내기 힘들기 때문에 현저한 인식 차이를 보이고 있다. 따라서 이 실험을 통해 향상된 인식결과를 얻기 위한 학습용 발음열 생성의 효용성을 확인할 수 있었다.

VI. 결론 및 향후과제

본 논문에서는 한국어의 음운변화 현상을 음소 변동 규칙과 변이음 규칙으로 설명하였다. 음소 변동 규칙은 13개의 필수 규칙과 4개의 수의 규칙으로 정의하고, 변이음 규칙은 3개를 정의하였다. 입력된 문자열 (기저형)에서 발음열 (표변형)의 생성은 필수 음소 변동 규칙, 수의적 음소 변동 규칙, 변이음 규칙들을 차례로 적용하여 생성하였다. 필수 음소 변동 규칙은 입력 문자열의 형태소와 그 주위의 문맥에 따라 4가지 테이블로 구성하여 선택적으로 적용하도록 하였고, 수의적 음소 변동 규칙은 비표준 발음열을 생성하는 규칙으로 다중 발음열을 생성하는데 사용되어졌으며, 규칙 적용으로 생성되지 않는 발음열은 예외 사전을 이용하도록 하였다.

발음열 자동 생성 모델은 본 논문에서 정의한 규칙을 이용하여 형태소 분석이 선행된 연결 단위의 입력을 받아 다중의 발음열을 출력하도록 구성되었다. 제한된 모델의 성능 평가를 위해서 전형적인 음운 변화를 포함한 실험 데이터와 PBS 낭독체 문장에 대한 실험 결과를 비교하였다.

본 시스템은 실험 자료에 대해 기존의 발음열 생성 시스템[9]보다 향상된 결과를 보여 주었다. 또한 형태음운론적 분석에 의해 구성된 시스템의 성능도 형태소 경계 정보와 접속정보의 사용여부에 따른 실험결과 높은 결과치를 얻었다. 또한 음성 인식에서의 발음열 생성기 사용

용도에 따른 실험을 수행하여 그 효용성을 평가하였다.

첫번째는 발음사전의 작성이다. 발음사전이란 각 단어에 대해 실제 인식의 기본 단위로 표기한 사전을 의미한다. 사전의 작성은 기존 학습 자료의 발화열들에서 사전 표제어 단위로 발화열들을 수집할 수도 있고, 문자사전에 대해 바로 발음열로 변화할 수도 있다. 앞선 방법에 대해 본 시스템을 이용하여 사전을 구축하였다. 이 경우 언절 내 표제어 간의 음운변화 현상이 반영된 발음형태가 기재된다. 실험 결과, 표제어 사이의 음운변화 현상을 반영한 발음사전의 경우가 반영되지 않은 경우보다 약 5~6% 정도 좋은 성능을 보였다.

두번째는 실제 음성 데이터가 표기된 학습용 발음열 생성에 관한 것이다. 발화열 작성에서 본 생성기는 모든 문장에 대해 일관적이고 다양한 발음열을 제공하므로 어떤 발화에 대해서도 적합한 발음열을 생성할 수 있다. 받아 적기로 작성된 발음열과 생성기 결과로 작성된 발음열을 각각 학습용 발음열로 사용하여 인식 실험을 수행한 결과 약 20%정도의 단어 인식을 향상시킬 수 있었다.

또한 발음열 생성기는 음성 합성에도 사용될 수 있다. 음성 합성은 무제한 단어의 합성이라는 점에서 실제 발음열을 발음사전에 수록할 수 없고, 실시간에 발음열을 생성해야 하므로, 정확하고 빠른 발음열 변환 시스템이 필요하다. 제안된 시스템은 이러한 상황에 적합하며, 특히 음성 합성에서 필수적인 요소인 언절 단위 발음열 생성을 만족시키는 조건을 갖추고 있어 합성용으로 확장 가능하다.

감사의 글

본 연구는 한국과학재단 목적기초연구 (과제번호 2000-1-30300-003-3) 지원으로 수행되었으며, 본 실험에 사용된 남독체 연속음성 DB는 삼성종합기술원의 지원으로 구축되었습니다.

참고 문헌

1. 배주재, 국어음운론 개설, 형설출판사, 1995.
2. 위선희, 정민화, "음운변화 규칙을 이용한 음성사전 생성", *HCI '97 학술대회 논문집*, 1997.
3. 이경님, 정민화, "외사형태소 단위의 연속 음성 인식", 제 15회 음성통신 및 신호처리 워크샵, pp. 309-314, 1998.
4. 이경님, 정민화, "한국어 연속음성인식을 위한 발음사전 구축", 제 27회 한국정보과학회 추계학술발표회, 2000.
5. 이성진, "Two-level 한국어 형태소 분석기", 한국과학기술원 전산학과 석사학위 논문, 1992.
6. 이호영, 국어음성학, 태학사, 1996.
7. 한국방송공사 편저, 표준한국어 발음 대사전, 어문각, 1993.
8. 한국의국어대학교, "한국어 특질에 관한 연구: 자동통역 전화시스템 구현을 위한 음운 및 문법구조 연구", 장기기초연구 보고서, 한국통신 연구개발단, 1994.
9. Jehun Jeon, Sunhwa Cha, Minhwa Chung, and Jun Park, "Automatic Generation of Korean Pronunciation Variants by Multistage Applications of Phonological Rules", *Proc. of*

International Conference on Spoken Language Processing, 1998.

10. Kimmo Koskenniemi, "A general computational model for word from recognition and production", *Proc. of the 8th International Conference on Computational Linguistics*, 1984.
11. Lori Lamel and Gilles Adda, "On Designing Pronunciation Lexicons for Large Vocabulary, Continuous Speech Recognition", *Proc. of International Conference on Spoken Language Processing*, 1996.
12. Maria-Barbara Wesenik, "Automatic Generation of German Pronunciation Variants", *Proc. of International Conference on Spoken Language Processing*, 1996.
13. Steve Young, Dan Kershaw, Julian Odell, Dave Ollason, Valtcho Valtcher, Phil Woodland, *The HTK Book (for HTK Verson 2.2)*, Entropic Cambridge Research Laboratory, 1999.

▲ 이 경 님 (Kyong Nim Lee)



1996년 2월: 명지대학교 컴퓨터공학과 (공학사)
 1998년 8월: 서강대학교 컴퓨터학과 대학원 (공학석사)
 1998년 9월~현재: 서강대학교 컴퓨터학과 대학원 박사과정
 ※ 주관심분야: 음성인식 및 언어처리, 음운 모델링, 발음사전

▲ 전 재 훈 (Je Hun Jeon)



1996년 2월: 서강대학교 전자계산학과 (공학사)
 1998년 2월: 서강대학교 컴퓨터학과 대학원 (공학석사)
 1998년 3월~현재: 삼성전자 대지텔 미디어총괄 컴퓨터시스템 사업부
 ※ 주관심분야: 음성인식, 발음열 생성

▲ 정 민 화 (Min Hwa Chung)



1984년 2월: 서울대학교 제어계측공학과 (공학사)
 1988년 5월: University of Southern California, 전기공학 (공학석사)
 1993년 8월: University of Southern California, 전기공학 (공학박사)
 1993년 12월~1994년 8월: 한국통신연구개발원 선임연구원
 1995년 9월~현재: 서강대학교 컴퓨터학과 부교수
 ※ 주관심분야: 음성인식 및 언어처리