

한국인 화자의 외래어 발음 변이 양상과 음절 기반 외래어 자소-음소 변환

Pronunciation Variation Patterns of Loanwords Produced by Korean and Grapheme-to-Phoneme Conversion Using Syllable-based Segmentation and Phonological Knowledge

류 혁 수¹⁾ · 나 민 수²⁾ · 정 민 화³⁾

Ryu, Hyuksu · Na, Minsu · Chung, Minhwa

ABSTRACT

This paper aims to analyze pronunciation variations of loanwords produced by Korean and improve the performance of pronunciation modeling of loanwords in Korean by using syllable-based segmentation and phonological knowledge. The loanword text corpus used for our experiment consists of 14.5k words extracted from the frequently used words in set-top box, music, and point-of-interest (POI) domains. At first, pronunciations of loanwords in Korean are obtained by manual transcriptions, which are used as target pronunciations. The target pronunciations are compared with the standard pronunciation using confusion matrices for analysis of pronunciation variation patterns of loanwords. Based on the confusion matrices, three salient pronunciation variations of loanwords are identified such as tensification of fricative [s] and derounding of rounded vowel [ɥi] and [wɛ]. In addition, a syllable-based segmentation method considering phonological knowledge is proposed for loanword pronunciation modeling. Performance of the baseline and the proposed method is measured using phone error rate (PER)/word error rate (WER) and F-score at various context spans. Experimental results show that the proposed method outperforms the baseline. We also observe that performance degrades when training and test sets come from different domains, which implies that loanword pronunciations are influenced by data domains. It is noteworthy that pronunciation modeling for loanwords is enhanced by reflecting phonological knowledge. The loanword pronunciation modeling in Korean proposed in this paper can be used for automatic speech recognition of application interface such as navigation systems and set-top boxes and for computer-assisted pronunciation training for Korean learners of English.

Keywords: loanwords, pronunciation variation, syllable-based, grapheme-to-phoneme, Korean

1. 서론

최근 음성인식이 폭넓게 사용됨에 따라, 내비게이션이나 TV의 셋톱박스과 같은 응용 인터페이스 환경으로 그 영역이 확장되고 있다. 이와 같은 응용 인터페이스 환경에서는 TV 프로그램명이나 내비게이션 응용에서 목적지 지명이나 상호명과

같은 POI(Point of Interest) 정보를 입력하기 위해 음성인식이 사용되고 있는데, 인식 대상이 되는 상호명이나 TV 프로그램명, 혹은 가수 이름이나 노래 제목 등에 외래어가 흔히 사용되고 있다. 외래어란 외국에서 유래한 말 중 한국어 음운, 문법, 의미 체계에 동화되고 일상생활에서 자주 사용됨으로써 한국어의 일부로 받아들여진 어휘를 말한다(차재은, 2007). 한국어의 경우, 표준국어대사전의 표제어의 약 4.7%만이 외래어인 반면(차재은, 2007), 방송 프로그램 제목의 63%가 외래어 어휘로 구성되어 있다고 보고된 바 있다(이은경, 1998).

응용 인터페이스에 대한 음성 인식 성능을 향상시키기 위해서는 외래어 발음이 음성인식기를 위한 발음 사전에 반영되어야 한다. POI나 멀티미디어에서는 주로 사전에 등재되지 않은 신조어의 비율이 높게 나타날 수밖에 없다. 만일 발음 사전을

1) 서울대학교, oster01@snu.ac.kr

2) 서울대학교, dix39@snu.ac.kr

3) 서울대학교, mchung@snu.ac.kr, 교신저자

이 논문은 SKT의 연구과제의 일환으로 수행되었습니다 (지원번호: 100-20150007).

접수일자: 2015년 8월 4일

수정일자: 2015년 9월 7일

게재결정: 2015년 9월 17일

수동으로 구성하게 될 경우, 새로운 신조어가 나타날 때마다 막대한 시간과 노동력을 들여서 지속적으로 업데이트를 해야 하는 문제가 있다. 따라서 한국어에서의 외래어 발음열 생성을 위해서는 외래어에 대한 자소 - 음소 변환기 (Grapheme-to-Phoneme, 이하 G2P)가 필수적이다.

한국어 발음열 생성에 관한 연구에 대해서는 그 동안 많은 연구가 이루어진 바 있으나(Kim et al., 2002; Yoon & Brew, 2006; Lee & Chung, 2007; Lee & Lee, 2009), 한국어에서의 외래어는 한국어 고유어와는 서로 다른 음운론적/음성학적 특성을 보인다(차재은, 2007; Kang, 1996; 류혁수, 정민화, 2014). 예를 들어, “SECRET”라는 단어는 <그림 1>에서 보이는 바와 같이 세 가지 발음 변이를 보일 수 있다. 치경 마찰음 /s/가 /s̰/로 경음화되는 현상은 일반적인 한국어 어휘에서는 일어나지 않고, 외래어에서만 발생한다(차재은, 2007; 최혜원, 2001). 한국어 발음열 생성에 관한 기존 연구에서도 외래어에 대해서는 발음열 생성 성능이 좋지 않음을 오류 분석을 통해 언급한 바 있다(Yoon & Brew, 2006; Lee & Lee, 2009). 따라서 응용 인터페이스에서의 외래어 음성인식을 위해서는 한국어 G2P 이외에도 한국어의 외래어에 대한 G2P가 필수적이다.

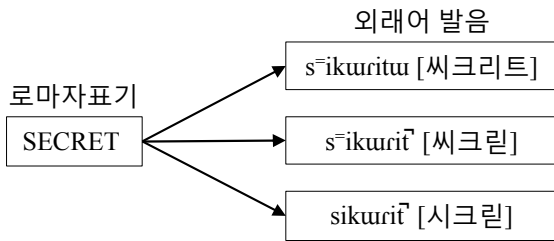


그림 1. 한국어의 외래어 발음 변이

Figure 1. Pronunciation variations of loanwords in Korean

한국어의 외래어에 대한 자소-음소 변환을 위해서는, 이에 대한 기초 연구로서 한국인 화자들의 외래어 발음 변이 양상에 대한 분석이 선행되어야 한다. 외래어의 발음 변이 양상에 대해 최혜원(2001)과 차재은(2007)과 같은 연구들이 있었다. 최혜원(2001)은 외래어 발음 실태에 대한 최초의 대규모 전면 조사로서, 표준 국어 대사전을 바탕으로 표기와 실제 발음의 차이가 크게 나타날 것으로 예측되는 어휘 170개를 선정하여 외래어 발음을 조사하였다. 하지만, 실태 조사가 목적이었기 때문에 현상 기술만을 하고 있을 뿐, 종합적 검토가 부재하다는 한계가 있다. 차재은(2007)의 연구에서는 국어사전 표제어로 등록된 외래어 가운데 표기와 발음이 다를 것으로 예상되는 어휘 104개를 선정하여 뉴스 방송 음성을 전사하였다. 실제 발음 전사 결과를 바탕으로 외래어 발음에 대한 필수 규칙과 수의적 규칙으로 구분하고, 외래어의 표준 발음정보를 선정하고자 하였다. 그러나 음성인식의 관점에서는 표준 발음 뿐 아니라 사용자의 다양한 발음 변이를 받아들일 수 있어야 하므로

표준 발음정보 선정에만 초점을 두고 있는 이 연구는 한계가 있다. 또한, TV 셋톱박스 혹은 내비게이션과 같은 응용 인터페이스에서는 새롭게 만들어진 외래어와 같이 사전에 등재되지 않은 어휘들이 많이 등장하지만, 국어 사전 표제어를 대상으로 한 소규모 어휘만을 이용한 발음 변이 조사로는 이러한 변이를 모두 담아내기에는 한계가 있다. 따라서 외래어 자소-음소 변환 연구에 앞서, 실제 응용 환경에서 사용되는 외래어를 대상으로 한국인 사용자들의 발음 변이 양상에 대한 분석이 진행될 필요가 있다. 이러한 필요에 의해 류혁수, 정민화(2014)의 연구에서 3,652개의 외래어 어휘를 대상으로 발음을 전사하여 기준 발음과 실제 발음 사이의 변이 양상을 분석하였다. 그 결과, 경음화, 과찰음화, 저모음화와 같은 주요 외래어 발음 변이를 관찰할 수 있었다. 류혁수, 정민화(2014)의 연구는 응용 어플리케이션 환경에서 많이 사용되는 대규모 외래어 어휘에 대해 발음 변이 양상을 분석하였다는 점에서 의의를 갖는다.

외래어에 대한 발음열 생성에 관해서도 그동안 다양한 연구들이 진행된 바 있다(Wang & Tong, 2014; Modipa & Davel, 2010). Wang & Tong(2014)의 연구는 음성 인식을 위해 중국 만다린에서의 외래어 발음모델링에 대해 다루었다. 영어와 중국어의 음운체계와 음절구조의 차이를 이용하여 영어 외래어의 발음사전을 확장시켰으며, 실험 결과, 확장된 발음사전을 이용하였을 때 외래어에 대한 음성인식성능이 개선됨을 확인하였다. Modipa & Davel(2010)은 남아프리카 공화국의 공용어 중 하나인 Sepedi어에서의 외래어 발음모델링에 관한 연구를 진행하였다. 영어와 다른 남아공 공용어에서 유래한 외래어 어휘의 발음열을 Sepedi어 자소 - 음소 변환 규칙을 이용하여 예측하였다. 그 결과 자소 - 음소 변환 규칙을 적용하여 발음열을 확장한 모델의 성능이 그렇지 않은 모델보다 우수한 것으로 나타났다. 하지만, 한국어의 외래어 발음열 자동 생성에 대해서는 그동안 많은 연구가 이루어지지 않았다. 이에 본 연구는 류혁수, 정민화(2014)의 후속 연구로서 외래어 어휘를 대상으로 한국인 화자의 외래어 발음 변이 양상을 분석하고, 이를 바탕으로 외래어 발음열을 자동으로 생성함으로써 자소-음소 변환 성능을 살펴보고자 한다. 본 연구에서 다루는 외래어는 로마자에 기반을 둔 외래어(로)로 한정한다.

2. 한국어에서의 외래어 발음 특성

2.1. 한국어와 영어의 음운론적 차이

한국어에서 사용되는 많은 외래어는 영어에서 유래한다. 따

4) 영어와 같이 로마자 기반의 외래어 뿐 아니라 중국어나 일본어와 같이 한자 문화권에서 유래한 외래어도 있으나, 지명이나 회사명과 같은 고유 명사가 대부분이고 그 수도 로마자 기반 외래어에 비해 많지 않기 때문에 본 연구에서는 다루지 않았다.

표 1. 외래어에서의 영어 음소의 대응
Table 1. Mapping between English and loanwords

범주	영어 음소	대응되는 외래어 발음	
		모음 앞	자음 앞/어말
마찰음	f, v, θ, ð, ʃ, ʒ	프(p), 브(b), 스(s), 드(d), 스(s), 즈(dʒ)	프(pu), 브(bu), 스(su), 드(du), 스(su), 지(dʒi)
파찰음	ts, dz	츠(tʃ), 즈(dʒ)	치(tʃi), 지(dʒi)
모음	ɔ, ə	오(o), 아(a)	

라서 기존 연구에서 언급된 외래어의 발음 특성을 알아보기에 앞서 본 장에서는 한국어와 영어사이의 음운론적 차이가 외래어 발음에 어떻게 작용하는지에 대해 간략히 정리한다.

한국어와 영어의 음운론적 차이는 크게 음소 체계에 의한 차이와 음절 구조에 의한 차이로 구분할 수 있다. 우선, 음소 체계에 의한 차이와 관련하여 영어의 음소 가운데 한국어 음소 체계에 존재하지 않는 음소들이 있다. 기본적으로 이들 음소들은 한국어의 외래어 표기법(국립국어원, 2015)에 의해 주변 환경에 따라 다음의 <표 1>과 같이 대응된다.

두 번째로, 한국어와 영어의 음절구조의 차이로 인해 한국어와 영어는 서로 다른 자음군, 음절말 구조를 갖는다. 자음군의 경우 한국어에서는 허용되지 않는 반면, 영어에서는 STRENGTHS /strenkθs/의 예에서 보는 바와 같이 음절초에서는 최대 3개의 자음까지, 음절말에서는 최대 4개의 자음이 허용된다. 또한 음절말 자음의 경우, 한국어에서는 7종성규칙에 의해 /ŋ(kʰ), ɫ(nʰ), ɰ(tʰ), ɽ(lʰ), ɱ(mʰ), ɸ(pʰ), ɷ(ŋ)/의 7개 자음만 허용되는 반면, 영어에서는 /h/를 제외한 모든 자음이 음절말에서 허용된다. 한국어의 외래어 발음에서는 모음 /u/를 삽입하는 방식으로 해결한다.

2.2. 한국어의 외래어 발음 변이 현상

비록 한국어에서의 외래어는 그 정의에 의해 한국어의 범주에 속하지만, 외래어 발음 변이는 한국어의 음소 변동 규칙과는 다른 특성을 보인다(Kang, 1996). 이러한 점으로 인해 외래어에서만 나타나는 특유의 발음 변이가 존재한다(차재은, 2007; 류혁수, 정민화, 2014). 기존 연구들에서는 외래어에서만 나타나는 발음 변이 현상으로 음절초 자음의 경음화, 마찰음

표 2. 기존 연구에서 제시된 외래어 발음 변이와 예시
Table 2. Pronunciation variations and the corresponding examples in loanwords presented in related studies

발음 변이	예시	
음절초 자음 경음화	GAME /k̚eim/	BOX /p̚ak̚s̚uw/
마찰음 /s/ 경음화	ACE /e̚is̚uw/	SIGN /s̚ain/
유음 /l/ 탈락	TELEVISION /t̚ɛ̚ɾ̚ɛ̚bidʒan̚/	PLASTIC /p̚h̚urasut̚h̚ik̚/
모음 변이	LIGHTER /lait̚a/	COLOR /k̚alla/

/s/의 경음화, 유음 /l/ 탈락, 모음 변이 등을 제시한 바 있다(최혜원, 2001; 차재은, 2007; 류혁수, 정민화, 2014). 각각의 변이와 예시는 다음의 <표 2>에 제시하였다.

3. 실험 방법

3.1. 코퍼스 및 전사

본 연구에서는 SK 텔레콤에서 제공한 로마자로 쓰인 외래어 단어 코퍼스를 사용하였다. 본 코퍼스는 텍스트 코퍼스로서, 셋톱박스, 음악, POI 영역에서 빈번하게 사용되는 14,449개 외래어 단어로 구성되어 있다. 외래어 코퍼스의 세부 구성 내역은 다음과 같다.

표 3. 외래어 단어 코퍼스
Table 3. Details of the Loanword Corpus

영역	단어 수	비율 (%)
셋톱박스	2,942	15.31
음악	2,212	20.36
POI	9,295	64.33
전체	14,449	100.00

3.2 발음 전사

외래어 단어 발음 전사에는 6명씩 5개 그룹으로 총 30명의 전사자가 참여하였다. 각 그룹은 2명의 서울대학교 언어학과 대학원생, 2명의 언어학과 학부생, 그리고 2명의 비언어학 전공자로 구성되어 있으며, 각 그룹의 전사자들은 영역에 무관하게 약 2,500 - 3,000개의 단어의 발음을 전사하였다. 언어학적 지식 여부에 따른 발음 차이를 반영하기 위해, 비언어학 전공자를 발음 전사에 포함시켰다.

전사자들에게 로마자로 쓰여진 외래어 단어들을 제공하고, 해당 단어들에 대해 자신들의 발음을 직접 전사하도록 하였다. 이 때, 영어와 같은 외국어를 사용하는 환경이 아니라 한국어로 대화하는 환경에서 가능한 모든 발음을 전사하도록 안내하였다. 이때, 한국어의 현실 발음을 반영하기 위해 /ㄱ/와 /ㄴ/, /ㅋ/와 /ㆁ/, /ㄷ/와 /ㄹ/ 및 /ㄷ/와 /ㄹ/은 구분하지 않고 각각 /ㄱ/, /ㅋ/, /ㄷ/로 대표하여 전사하도록 하였다. 한 단어에 6명의 전사자가 발음을 전사하고 있기 때문에 전사자들의 발음열을 모두 채택하면 지나치게 많은 발음열이 발생할 수 있다. 또한 전사자의 방언이나 개인 특성으로 인한 발음 변이가 발생할 수 있기 때문에, 이를 방지하기 위하여 3명 이상의 전사자들이 공통으로 전사한 발음열을 외래어 단어의 발음으로 채택하였으며, 그 결과 16,968개의 발음열이 채택되었다.

3.3 기준 표기

외래어의 발음변이 양상을 분석하기 위해서는 실제 발음과

기준 발음을 비교하는 작업이 필요하다. 류혁수, 정민화(2014)의 연구에서는 로마자로 된 외래어를 표준 외래어 표기법(국립국어원, 2015)에 의해 한글 표기로 변환한 후, 해당 표기에 따른 표준 발음을 기준 발음으로 간주하였다. 하지만, 연규동(2006)이나 정국(2003)의 연구 등에서 지적하는 바와 같이 표준 외래어 표기법이 한국어 화자들이 사용하는 외래어 표기와 괴리를 보이고 있기 때문에 비교 대상의 기준으로 적절하지 않다는 문제가 있다. 예를 들어, LOBSTER의 경우 외래어 표기법에 따르면 ‘로브스터’가 되어야 하나, 이와 같이 표기하기 보다는 한국어 화자들은 ‘랍스터’라는 표기를 선호한다.

이와 같은 이유로 본 연구에서는 발음변이 양상 분석에서의 기준을 정하기 위하여 한국어 화자들이 실제 사용하는 표기를 기준으로 삼고자 하였다. 이를 위해 2명씩 1개 그룹으로 총 8명의 전사자에게 한국어로 글을 쓰는 환경에서 사용가능한 모든 표기를 전사하도록 안내하였다. 표기전사자들은 모두 서울대학교 언어학과 대학원생으로 구성되었다. 전사자들이 전사한 표기를 가운데 가장 많은 전사자들이 선택한 표기를 대표 표기로 선정하였으며, 만일 동수가 나온 경우에는 연구자가 가장 적절하다고 생각되는 표기를 채택하였다.

3.4 발음변이 양상 분석

발음변이 양상을 분석하기 위해 기준 발음과 실제 발음을 정의하고 이를 phone-like unit (PLU)로 변환하였다. 먼저, 기준 발음은 3.3에서 도출한 기준 표기에 대해 서울대학교 음성언어 처리연구실의 G2P(Lee & Chung, 2007)를 이용하여 기준 발음열을 생성하였다. 실제 발음은 3.2의 발음 전사를 통해 채택된 발음열로 정의하였다. 위에서 정의한 기준 발음열과 실제 발음열을 정렬하여 혼동행렬을 구성하고, 200건 이상 발생한 음소들에 대해 자/모음의 평균 변이 비율보다 높게 나타나는 변이를 주요 발음 변이로 도출하였다. 혼동 행렬의 각 행에는 기준 발음이 제시되어 있으며, 각 열은 전사자들의 실제 발음을 의미한다.

3.5 자소-음소 변환

3.5.1 음절 기반 분절 정보

한국어에서의 외래어는 동일한 자소 형태를 가지고 있다고 할지라도 다양한 발음열을 가지고 있다. 예를 들어, 외래어인 “EMMA”는 /ɛma/ 또는 /ɛm̃ma/로 실현될 수 있다. 자소열이 완전히 동일하기 때문에 자소열의 정보만으로는 위의 예에서 보는 바와 같이 서로 다른 발음열을 생성하기 어렵다는 문제가 있다.

2.1에서 언급한 바와 같이, 한국어와 영어에서 허용하는 음절 구조에서 차이가 발생하기 때문에 외래어가 되면서 음절의 분절이 서로 다르게 실현될 수 있다. <그림 2>의 예시에서 보는 바와 같이 EMMA /ɛma/의 경우 자소 MM은 두 번째 분절로 포함시키는 것이 타당하다. 첫 번째 발음 분절에는 /ɛ/ 이기

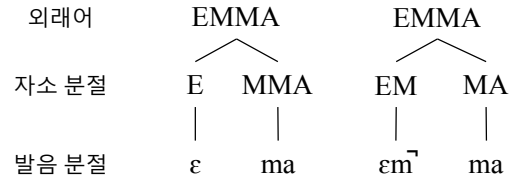


그림 2. 분절 예시 - EMMA
Figure 2. Segmentation example - EMMA

때문에 자소에서는 E를 발음 분절 /ɛ/로 사상되어야 하기 때문이다. 반면, 발음열이 /ɛm̃ma/일 때는 자소의 MM이 분리되어 각각의 분절로 할당된다.

이와 같이 동일한 자소열을 갖더라도 발음열의 음절 구조에 따라 자소열이 서로 다른 분절 구조를 가지게 됨을 알 수 있다. 다시 말해, 외래어 단어 뿐만 아니라 외래어의 분절 정보가 함께 주어진다면 주어진 단어에 대한 발음열을 보다 효과적으로 예측할 수 있을 것으로 기대할 수 있다. 예를 들어 EMMA라는 외래어 단어가 주어졌을 때, E | MMA 라는 분절 구조가 나타난다면 해당 외래어의 발음열은 /ɛma/로 생성할 수 있고, 반면 EM | MA 라는 분절 구조가 나타난다면 /ɛm̃ma/라는 발음열을 예측할 수 있을 것이다. 따라서 본 연구에서는 한국어에서 사용하는 외래어의 다양한 발음열을 생성하기 위해 음절 기반의 분절 정보를 이용하는 자소-음소 변환 방법을 제안하고자 한다.

본 연구에서 제안하는 분절 규칙은 다음과 같다. 기본적으로 외래어는 로마자의 음절에 기반하여 분절을 나눈다. 하지만 2.1에서 언급한 바와 같이 한국어와 영어는 허용하는 음절 구조에서 차이가 있으므로 분절 구분을 위해 자음군과 음절말 자음에서의 모음 삽입을 고려하여야 한다. 따라서 삽입된 모음에 대해서도 분절을 나눈다. 예를 들어 SECRET은 영어에서는 2음절로 구성되어 있으나, 한국어에서 외래어로 실현될 경우 <그림 3>에서 보는 바와 같이 자음군과 음절말 자음에서 모음 삽입이 발생하기 때문에 3음절 혹은 4음절로 분절이 나누어지게 된다.

3.5.2 혼련 방법

외래어 자소-음소 변환을 위해서 n-gram 결합 연쇄 모델(n-gram joint sequence model)을 사용하는 Sequitur G2P(Bisani

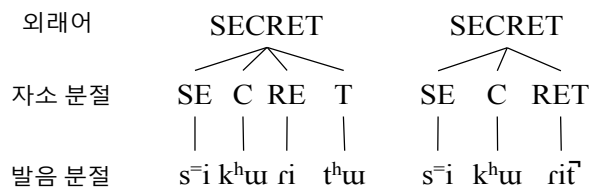


그림 3. 모음 삽입을 고려한 분절 - SECRET
Figure 3. Segmentation considering vowel epenthesis - SECRET

표 4. 분절 및 발음 사전
Table 4. Segment and Pronunciation dictionaries

사전 종류	외래어 단어	분절/발음
분절 사전	SECRET	SE C RE T
발음 사전	SECRET	s ^h i k ^h u r i t ^h u

& Ney, 2008)를 적용하였다. 결합 연쇄 모델은 자소의 연쇄와 그에 해당하는 음소의 연쇄를 고려하여 그 연쇄의 확률을 계산하는데, 이러한 모델을 사용하는 Sequitur G2P는 자소 - 음소 변환에서 최신기법의 하나이다(Jouvet et al., 2012; Schlippe et al., 2012; Schlippe et al., 2014).

외래어 자소-음소 변환 실험은 앞서 3.5.1에서 제안한 음절 기반 분절 정보를 이용한 방법과 베이스라인을 비교하는 방식으로 수행되었다. 베이스라인은 분절 정보를 제외한 외래어의 로마자 자소열과 발음열을 이용하여 G2P를 훈련시켰다. 본 연구에서 제안한 방법으로 음절 기반의 분절 정보를 이용하기 위해 2개의 발음 사전을 준비하였다. <표 4>에서 보는 바와 같이 하나는 분절 사전으로서 외래어 로마자 자소의 분절 구성에 대한 정보를 담고 있으며, 다른 하나는 발음열 사전으로 외래어 로마자 자소에 대한 발음열 정보를 담고 있다. 이렇게 구성된 분절 사전의 분절 정보와 발음 사전의 발음을 사상(mapping)하여 훈련을 진행한다.

자소-음소 변환기의 훈련은 반복을 통해 수행된다. 처음 훈련할 때는 매우 간단한 모델을 생성하지만, 훈련 횟수가 늘어남에 따라 이전에 훈련된 모델을 이용하여 또 다시 훈련한다. 본 실험에서는 총 7회 반복하여 훈련하였다. 각 반복이 완료될 때마다 훈련된 모델에 테스트 데이터를 적용하여 성능을 측정하였다.

음성 인식의 관점에서 사용자의 다양한 발음 변이를 인식할 수 있어야 하므로, G2P는 다중 발음열을 생성할 수 있어야 한다(Jouvet et al., 2012). 이러한 이유로 본 연구에서 제안하는 자소-음소 변환에서도 다중 발음열을 생성할 수 있도록 하였다. 이를 위해 발음열의 계산모델 확률 누적합이 0.5를 초과할 때까지 발음열을 생성하도록 하였다. 하지만, 개별 발음열의 확률이 낮을 경우, 확률 누적합을 충족하기 위해서 타당성이 떨어지는 발음열이 지나치게 많이 생성되는 문제가 있으므로, 생성되는 발음열의 개수는 10개로 제한하였다.

또한 좌우 문맥 정보를 함께 제공하여 외래어 발음을 생성하고자 하였다. 이를 위해 3-그램 (좌우 1개 음소 문맥)부터 9-그램 (좌우 4개 음소 문맥)까지 총 4가지 문맥 환경에 대해 실험을 진행하였다.

5) 본 연구에서는 G2P의 입력으로 로마자 표기를 사용하고 있다. 이는 네비게이션이나 TV 프로그램 등에서 외래어가 로마자로 표기되어 있는 경우도 많기 때문에 이에 대한 음성 인식을 위해서 G2P를 사용하는 경우를 상정하였기 때문이다. 외래어 단어를 한글로 표기하는 다른 응용에서는 한글로 입력된 외래어의 발음열 생성에 대한 대책이 필요하나, 본 연구의 범위를 벗어나기 때문에 후속 연구에서 다루고자 한다.

3.5.3 성능 평가

자소-음소 변환 성능을 평가하기 위해서 2가지 종류의 단위를 사용하였다. 하나는 음소 오류 비율(Phone error rate, 이하 PER)/단어 오류 비율(Word error rate, 이하 WER)이고 다른 하나는 정밀도(precision)/재현율(recall)/F-score이다. PER과 WER은 발음열 생성 결과에 대한 평가를 위해 흔히 사용되는 측정 단위이고(Schlippe et al., 2014; Hahn et al., 2012), 정밀도와 재현율 역시 다중 발음열의 질을 평가하기 위한 지표로 사용된다(Jouvet et al., 2012; Bang et al., 2014). F-score는 정밀도와 재현율의 조화 평균으로 계산한다.

3.5.4 실험 구성

3.5.2에서 제시된 훈련 방법을 이용하여, 본 연구에서는 1) 제안한 분절 정보를 이용한 방법과 베이스라인의 성능 비교와 2) 외래어 영역에 따른 자소-음소 변환 성능 차이에 대한 비교의 두 가지 실험을 시행하였다. 첫 번째 실험에서는 외래어 어휘 영역과 상관없이 훈련 집합과 테스트 집합을 구분하였다. 테스트 집합 특성으로 인한 성능 변화를 방지하기 위해서 10회 교차 검증(10-fold cross validation)을 시행하였다. 또한 과라미터 최적화를 위해 훈련 집합의 5%를 개발 집합(development set)으로 이용하였다. 실험 데이터의 세부 정보는 다음의 <표 5>와 같다.

표 5. 베이스라인과 제안 방법 비교 실험 세부 정보
Table 5. Statistics of the data for the comparison between the baseline and the proposed method

	단어 수	비율 (%)
훈련 집합	12,350	85.50
개발 집합	650	4.50
테스트 집합	1,445	10.00
총합	14,445	100

두 번째 실험은 외래어 어휘의 영역이 G2P 성능에 미치는 영향을 알아보기 위한 목적으로 시행하였다. 본 실험에서 외래어 어휘는 POI, 셋톱박스, 음악 영역으로 구분된다. 3.1의 <표 3>에서 보는 바와 같이 셋톱박스와 음악 영역의 데이터는 훈련에 이용하기에 너무 적은 관계로, POI 데이터만을 이용하여 G2P를 훈련하였다. 테스트 집합으로는 세 종류의 집합을 모두 사용하였다. 세부 정보는 다음의 <표 6>과 같다.

표 6. 어휘 영역에 따른 성능 비교 실험 세부 정보
Table 6. Statistics of the data for comparison of performance in terms of domains

	영역	단어 수	비율 (%)
훈련 집합	POI	7,600	63.96
개발 집합	POI	400	3.37
테스트 집합	셋톱박스	1,295	10.89
	음악	1,295	10.89
	POI	1,295	10.89
총합		11,885	100

표 7. 한국인 외래어 발화의 모음 혼동 행렬
Table 7. Confusion matrix for vowels of loanwords produced by Korean speakers

기준 발음	실현 발음																	
	(%)	ㅏ	ㅑ	ㅓ	ㅕ	ㅡ	ㅣ	ㅘ	ㅙ	ㅛ	ㅜ	ㅠ	ㅚ	ㅜ	ㅝ	ㅞ	ㅟ	
ㅏ (7,358)	94.58							2.08										1.23
ㅑ (4,081)	2.16	92.68	2.21															
ㅓ (5,741)	1.15	1.39	96.76															
ㅕ (2,285)		1.10		92.95								2.37						2.28
ㅡ (11,252)					98.40													
ㅣ (12,079)						97.63												
ㅘ (6,887)	1.49							93.95										2.11
ㅙ (270)	1.11								95.19			2.22						
ㅛ (328)										98.17								
ㅜ (140)									2.14		97.14							
ㅠ (579)				4.50								91.87						1.21
ㅚ (73)						1.37	2.74			2.74		1.37	83.56				5.48	2.74
ㅜ (186)	2.15													93.55	1.08			2.15
ㅝ (288)				1.39										1.74	95.82			
ㅞ (585)							21.03										78.12	
ㅟ (288)													7.77				1.46	87.38

* 1% 미만의 변이는 표시하지 않음.
* 괄호는 해당 음소의 발생 건수를 의미함.

표 8. 한국인 외래어 발화의 자음 혼동 행렬
Table 8. Confusion matrix for consonants of loanwords produced by Korean speakers

가. 초성 자음

기준 발음	실현 발음																			
	(%)	ㅂ	ㅃ	ㅍ	ㅌ	ㅍ	ㅌ	ㄱ	ㅋ	ㅋ	ㅅ	ㅆ	ㅅ	ㅈ	ㅊ	ㅈ	ㅊ	ㅌ	ㅎ	ㅏ
ㅂ (3,654)	97.40	1.48																		
ㅃ (45)	33.33	62.22	4.44																	
ㅍ (3,978)			98.36																	
ㅌ (3,095)				98.58																
ㅍ (76)				15.79	55.26	23.68								5.26						
ㅌ (4,465)						95.56														3.18
ㄱ (1,347)							94.05	1.56	1.19											2.16
ㅋ (44)							11.36	61.36	27.27											
ㅋ (3,885)								3.35	93.79											2.50
ㅅ (2,270)										89.77	1.81				6.83					
ㅆ (47)										23.40	70.21	4.26			2.13					
ㅅ (734)											1.09	93.99			2.46					
ㅈ (6,396)										6.43				38.75	54.00					
ㅊ (895)					2.13									2.01	94.52					
ㅈ (1,195)																97.50				1.04
ㅊ (4,616)																	97.03			1.21
ㅌ (2,417)																			99.34	
ㅎ (2,108)																	1.85			96.34

* 1% 미만의 변이는 표시하지 않음.
* 괄호는 해당 음소의 발생 건수를 의미함.

나. 종성 자음

기준 발음	실현 발음								
	(%)	ㅂ	ㄷ	ㄱ	ㄹ	ㅁ	ㄴ	ㅇ	삭제
ㅂ (570)	99.64								
ㄷ (483)		91.67							
ㄱ (1,261)			99.26						
ㄹ (5,906)				97.07					
ㅁ (3,814)					99.19				
ㅁ (1,238)						98.58			
ㅇ (1,038)							1.16	3.00	95.55

* 1% 미만의 변이는 표시하지 않음.
* 행의 합이 100%가 되지 않는 경우는 해당 종성 자음 뒤에 모음 삽입이 일어나 초성 자음으로 실현된 경우임.
* 괄호는 해당 음소의 발생 건수를 의미함.

4. 실험 결과

4.1 외래어 발음변이 양상 분석

한국어 화자의 외래어 실제 발음과 기준 발음을 비교하여 외래어 발음 변이 양상을 분석한 결과, 14,449개 단어에 대해 전체 107,879개 음소로 구성되어 있었으며 모음이 52,338개, 자음이 55,541개의 음소로 구성되어 있었다. 이를 기반으로 혼동 행렬을 구성한 결과는 <표 7>, <표 8>과 같았다. 혼동 행렬 분석 결과, 외래어 발음 변이 비율은 다음과 같이 나타났다. <표 9>에서 볼 수 있는 바와 같이 기준 발음열 대비 발음 변이 비율은 전체 음소에 대해 7.84% 이었고, 모음에서는 4.97%, 자음에서는 10.54%의 변이 비율을 보였다. 자음의 변이 비율이 모음의 변이 비율보다 2배 이상 높게 나타남을 볼 수 있었다. 이는 자음의 전체 변이인 5,859건인 반면, /s/의 경음화가 총 3,452건 발생하여 자음 변이의 절반 이상을 차지하기 때문인 것으로 나타났다. 따라서 /s/의 경음화가 자음 변이 비율을 모음 변이 비율보다 훨씬 높게 나타나게 하는 원인으로 작용함을 판단할 수 있다.

표 9. 외래어 발음 변이 비율

Table 9. Variation rate of pronunciation in loanwords

	전체 음소 #	변이 음소 #	변이 비율
전체	107,879	8,458	7.84%
모음	52,338	2,599	4.97%
자음	55,541	5,859	10.54%

<표 7>과 <표 8>에서 제시한 혼동 행렬에 기반하여, 3.4에서 언급한 기준 (200건 이상 발생한 음소 & 자/모음 평균 변이 비율 [10.54%/4.97%] 이상의 변이)에 따라 주요 발음 변이를 추출하였다. 주요 발음 변이 분석 결과는 다음의 <표 10>과 같이 마찰음 /s/의 경음화, 원순모음 /qi/와 /we/의 원순성 약화로 나타났다.

표 10. 주요 발음 변이 양상

Table 10. Patterns of salient pronunciation variation

기준 발음	실현 발음	빈도	비율 (%)	분류
ㅅ /s/	ㅆ /s̺/	3,452	54.00	경음화
기 /qi/	이 /i/	123	21.03	원순성 약화
게 /we/	계 /je/	16	7.77	

먼저, 마찰음 /s/의 경음화는 기존 연구에서도 많이 언급된 바 있으며, 가장 대표적인 외래어에서의 발음 변이 양상이라고 할 수 있다. 치경 마찰음 /s/가운데 절반 이상이 경음화가 이루어지고 있음을 알 수 있다.

또한, 주요 발음 변이 양상이 나타나는 환경의 특성을 알아보기 위해 /s/가 /s̺/로 변화하는 후행하는 모음의 비율을 조사

하였다. <표 11>은 /s/의 경음화가 일어나는 변이의 후행 환경을 고빈도 순으로 상위 5건을 보여주고 있다. <표 11>에서 볼 수 있는 바와 같이, /s/의 경음화는 모음 /u/ 앞에서 절반을 상회하는 가장 높은 비율로 나타나고 있다. 또한 비율은 다소 낮지만, 다양한 모음 앞에서 /s/가 경음화되고 있음을 알 수 있다.

표 11. /s/ 경음화에서의 후행 모음 발생 빈도, 비율 및 예시
Table 11. Frequency, proportion, and examples of the following vowels in /s/ tensification

후행 모음	빈도	비율(%)	예시
u	2,052	59.44	CHRISTMAS /kʰurɨs̺tʰmas̺u/
i	418	12.11	AGENCY /eɪdʒən̺s̺i/
ɛ	313	9.07	ANDERSEN /an̺tɛrus̺ɛn̺/
o	191	5.53	GLOSOFT /kʰulos̺ɔp̺hʊt̺u/
a	153	4.43	INSIDER /in̺s̺aɪt̺a/

두 번째 주요 발음 변이 양상은 원순모음 /qi/와 /we/의 원순성 약화를 들 수 있다. 이들 현상은 원순 모음인 기/qi/와 게/we/가 원순성을 잃고 각각 평순 모음인 이/i/와 계/je/로 변이가 이루어지고 있다. 원순성 약화 현상이 발생하는 환경을 알아보기 위하여 선행하는 초성 자음의 비율을 분석한 결과, 다음의 <표 12>와 같이 나타났다. <표 12>에서 보는 바와 같이, 원순성 약화 현상은 치경 마찰음 /s/가 선행하고 있을 때 거의 대부분 나타나는 것을 관찰할 수 있었다. 기/qi/의 경우에는 <표 12>의 CLASH나 FISH의 사례에서 보는 바와 같이 주로 어말에서 실현되고 있는 반면, 게/we/의 경우에는 CHEF나 MICHELLE의 사례에서 보는 바와 같이 주로 어두나 어중에서 실현되고 있다는 차이를 보였다. 또한, 기/qi/에서는 FISH /p̺h̺is̺i/의 사례에서 볼 수 있는 바와 같이 원순성 약화 변이가 발생할 때, /s/의 경음화를 동반하기도 하는 반면, 게/we/의 계/je/로의 원순성 약화에서는 /s/의 경음화를 동반하지 않는다는 특징을 관찰할 수 있었다.

표 12. 원순성 약화에서의 선행 자음 발생 빈도, 비율 및 예시
Table 12. Frequency, proportion, and examples of the preceding consonants in derounding

변이 음소	선행 자음	빈도	비율(%)	예시
qi	s	118	95.93	CLASH /kʰulles̺i/ FISH /p̺h̺is̺i/
we	s	16	100	CHEF /s̺jep̺h̺u/ MICHELLE /mis̺jel/

4.2 외래어 자소-음소 변환

4.2.1. 베이스라인과 제안 방법의 비교

<그림 4>와 <그림 5>는 베이스라인과 제안한 방법의 PER과 WER 비교 실험 결과를 보여 준다. 그림에서 볼 수 있는 바와 같이 3-그램에서 9-그램까지 좌우 문맥을 고려하여 변환한 결

과, 3-그램을 제외한 모든 문맥 범위에서 제안한 방법의 PER 과 WER 값이 베이스라인보다 우수하게 나타나는 것을 볼 수 있었다. 특히 좌우 3개씩의 음소 문맥을 고려한 7-그램에서 가장 낮은 PER과 WER을 보였다. 제안한 방법 가운데 가장 성능이 좋았던 7-그램에서는 34.95%와 9.01%의 WER과 PER을 보인 반면, 베이스라인 가운데 가장 성능이 좋았던 5-그램에서는 37.21%와 9.50%의 WER과 PER을 보였다. 이는 제안한 방법이 WER과 PER에서 각각 6.07%와 5.16%의 상대향상율을 보임을 의미한다. 한편, 훈련이 진행됨에 따라 약 3회 정도 반복되었을 때, 성능이 포화상태에 도달하는 것을 관찰할 수 있다.

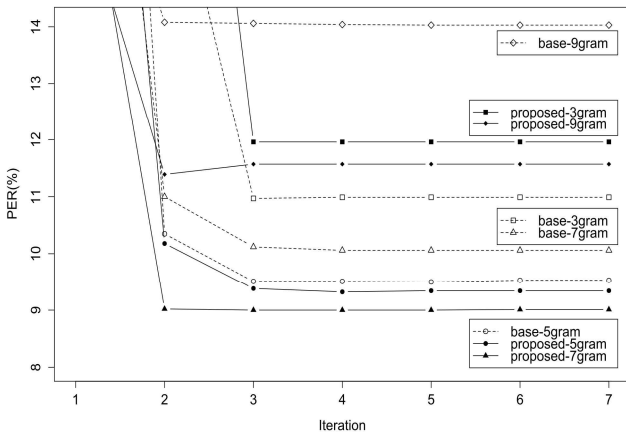


그림 4. 베이스라인과 제안 방법의 음소 오류 비율 결과 비교
Figure 4. Comparison using phone error rate between the baseline and the proposed method

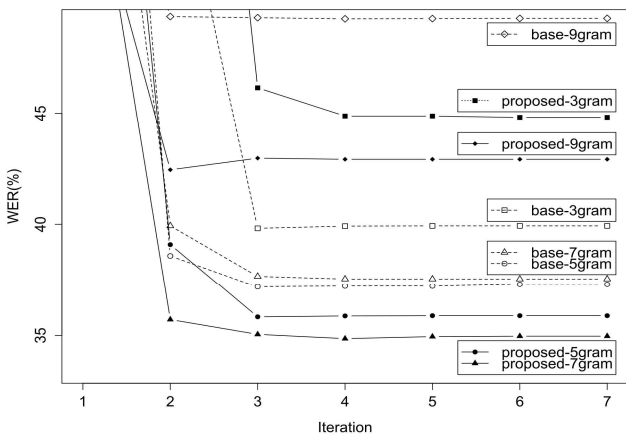


그림 5. 베이스라인과 제안 방법의 단어 오류 비율 결과 비교
Figure 5. Comparison using word error rate between the baseline and the proposed method

성능 평가를 위해 사용한 두 번째 측정 기준은 precision과 recall을 반영한 F-score이다. F-score에 의한 성능은 다음의 <그림 6>과 같다. 실험 결과, 앞서 PER과 WER을 이용한 성능 평가 결과와 동일하게 제안한 방법의 성능이 베이스라인의 성능

보다 우수하게 나타나고 있다. 제안한 방법 가운데 7-그램 환경에서 F-score가 0.570으로 가장 우수하게 나타난 반면, 베이스라인에서는 5-그램이 가장 높은 0.553의 F-score를 보였으며, 제안한 방법이 베이스라인 대비 3.07%의 상대향상율을 보인다.

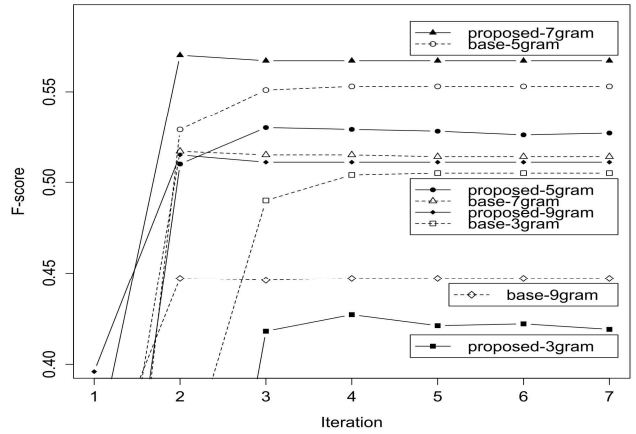


그림 6. 베이스라인과 제안 방법의 F-score 결과 비교
Figure 6. Comparison using F-score between the baseline and the proposed method

4.2.2. 어휘 영역에 따른 성능 비교

앞서 3.5.2에서 언급한 바와 같이, 본 연구에서 사용된 어휘는 음악, POI, 셋톱박스 영역으로 구분할 수 있다. 본 실험에서는 어휘 영역의 차이가 자소-음소 변환 성능에 영향을 미치는지 여부를 관찰하기 위해 어휘 영역에 따른 자소-음소 변환 성능을 비교하였다. 3.5.2에서 제안한 방법론에 따라 훈련하였으며, 4.2.1에서 가장 좋은 성능을 보인 7-그램의 문맥 범위를 적용하여 성능을 측정하였다. 실험 결과는 아래의 <표 13>에서 볼 수 있다.

표 13. 어휘 영역에 따른 G2P 성능 비교
Table 13. Experimental results in terms of domains

훈련 집합	테스트 집합	PER(%)	WER(%)	F-score
POI	POI	10.01	37.68	0.553
	음악	13.85	43.36	0.478
	셋톱박스	15.46	52.50	0.376

<표 13>에 따르면, 훈련 집합의 어휘 영역과 다른 어휘 영역으로 구성된 테스트 집합으로 성능을 평가하였을 때, 동일한 어휘 영역으로 구성된 테스트 집합보다 성능이 저하되는 것을 관찰할 수 있었다. 이를 통해, 본 연구에서 제안한 한국어 자소-음소 변환에서 외래어 어휘 영역의 차이가 성능에 영향을 미치고 있음을 알 수 있다.

4.3 G2P 오류 분석

4.2.1에서 살펴 보는 바와 같이, 비록 본 연구에서 제안한 자

소-음소 변환 성능이 베이스라인보다 높게 나타났으나, 일부 발음열 생성 결과에서 정답 발음열과 일치하지 않는 문제가 있었다. 향후 G2P 성능 개선에 도움이 되기 위해서 본 절에서는 발음 생성 결과와 정답 발음열이 일치하지 않는 현상을 범주별로 분석하여 <표 14>에서 제시하였다.

표 14. G2P 오류 범주 구분 및 예시
Table 14. Error classification and examples of G2P

범주	외래어 어휘	정답 발음열	생성 발음열
모음 교체	CAJUN	k ^h eiɟɟun ^ˀ	k ^h adɟun
	ATMACE	et ^ˀ meis ^ˀ u	et ^ˀ mas ^ˀ u
	CENTRAL	s ^ˀ en ^ˀ tʰural	s ^ˀ en ^ˀ tʰural
	EGO	iko	ɛko
	ETERNEL	it ^h ʌnal	ɛt ^h ʌnɛl
	EVENT	ipen ^ˀ tʰu	ipum ^ˀ tʰu
영어 이외의 외래어 발음	DEBUT	teɟju	teɟat ^ˀ
	CHALET	sjalle	ʃʌllet ^ˀ
	COREE	k ^ˀ ore	k ^ˀ ori
	CHI	k ^h ai	ʃi
목음	COMB	k ^h ʌm ^ˀ	k ^h om ^ˀ pu
	DRINKS	tʰurɪŋs ^ˀ u	tʰurɪŋk ^h us ^ˀ u
어중 /s/의 경음화	PANASONIC	p ^h anasonik ^ˀ	p ^h anas ^ˀ onik ^ˀ
	BESETO	pɛset ^h o	pɛs ^ˀ ɛt ^h o
	ASAKAN	asak ^h an ^ˀ	as ^ˀ ak ^h an ^ˀ
경음/격음 교체	BIZCAFE	pizuk ^h ap ^h ɛ	pizuk ^ˀ ap ^h ɛ
	BOUTIQUE	put ^h ik ^h u	put ^ˀ ik ^ˀ u

발음열 생성 오류 가운데 가장 빈번하게 나타나는 현상은 모음 교체 오류로 나타났다. 이는 동일한 모음 자소에 대해 2개 이상의 음가가 가능할 경우, G2P가 이를 제대로 구분하지 못함으로써 발생하는 오류이다. 예를 들어 모음 자소 A와 E에 대해 각각 /ei/, /ʌ/, /a/와 /e/, /i/가 가능한데, G2P가 이를 제대로 구분해 주지 못함으로써 오류가 발생하였다. <표 14>에서 볼 수 있는 바와 같이 대부분의 모음 교체 오류는 A와 E를 각각 /a/와 /e/로 실현시키기 때문에 오류로 나타나는 것을 관찰할 수 있다. 이는 본 연구에서 사용한 어휘 특성상 영어 단어 이외에 신조어와 고유어가 많이 포함되어 있어서 훈련데이터에서 A와 E가 철자 발음 그대로 각각 /a/와 /e/로 실현되는 사례가 많았기 때문으로 추정된다. 단, <표 14>의 EVENT의 경우 첫 번째 E는 발음을 높게 생성한 반면, 두 번째 E에서 /u/를 생성하는 것을 볼 수 있다. 이는 훈련 데이터에서 유사한 음소 환경을 갖는 EVEN /ipum^ˀ/이 있어서 이를 학습하였기 때문으로 추정된다. 이러한 오류는 G2P의 훈련 데이터 규모를

늘림으로써 상당 부분 줄일 수 있을 것으로 기대된다.

두 번째 사례는 영어 이외의 다른 언어에서 유래한 외래어의 경우 일부 발음열 생성에서 오류를 보였다. <표 14>의 예에서 볼 수 있는 바와 같이, 프랑스어 혹은 그리스어 등에서 유래한 외래어의 경우 발음 규칙이 영어와 상이하기 때문에 어말 자음의 탈락 (DEBUT, CHALET)이나 자음 음가 (COREE, CHI)를 제대로 예측할 수 없다는 한계가 있었다. 이와 같은 어휘의 경우에는 좌우 음소 문맥과 분절 정보만으로는 G2P에서 처리하기에 어려움이 있기 때문에 예외 사전과 같은 전처리)를 통해서 해결할 수 있을 것으로 보인다.

또한 일부 외래어의 경우 자소는 존재하지만 목음으로 발음되는 어휘들이 있었다. COMB나 DRINKS의 예에서 볼 수 있는 바와 같이 각각 자소 B와 K가 존재하지만 실제 발음에서는 실현되지 않는 것으로 발음 전사가 이루어 졌다. 하지만, 이러한 예가 상당히 적기 때문에 이에 대해 G2P가 제대로 훈련되지 못한 관계로 오류가 발생한 것으로 보인다. 이는 훈련 데이터가 늘어난다면 오류가 감소할 것으로 생각된다. 또한, 다른 한편으로는 화자의 어휘지식이나 어휘가 사용되는 환경에 따라 이를 목음으로 발화하지 않는 경우도 가능하기 때문에, 음성 인식이라는 관점에서 살펴보면 목음으로 처리되지 않은 발음열도 필요할 것으로 보인다. 따라서 목음이 반영되지 않은 발음열을 반드시 오류라고 간주할 수만은 없다.

마지막으로 정답 발음열과 생성된 발음열이 일치하지 않는 범주로는 어중 /s/의 경음화와 경음과 격음의 교체 현상이 있었다. 어중 /s/의 경음화의 경우, 차재은(2007)에서 언급한 바와 같이 외래어 발음에서 주의적으로 나타나는 현상이다. 음성 인식 관점에서는 발화자의 가능한 발음을 가급적 모두 반영해 줄 필요가 있음을 고려할 때, 발음열 생성 결과가 정답 발음열과 차이를 보인다고 해서 반드시 자소-음소 변환 결과가 잘못된 것으로 간주할 필요는 없을 것으로 생각된다. 또한 경음과 격음의 교체 현상의 경우에서도 <표 14>의 BIZCAFE와 BOUTIQUE의 사례에서 볼 수 있는 바와 같이 발화자에 따라서는 이를 격음이 아닌 경음으로 발음할 가능성도 높다. 이러한 점을 고려할 때 생성된 발음열의 결과 역시 오류라고만 보기에는 어려움이 있다.

6) 자소 E의 경우, ICE /ais^ˀu/ 와 같이 /u/로 실현될 수도 있으나, G2P에서 오류로 나타나지 않았기 때문에 설명에서 제외하였다.

7) 영어사전에 등재되어 있는 단어의 경우 외래어 G2P 처리 이전에 영어 발음 사전을 검색하여 발음열을 생성하고 미등록어에 대해 G2P로 처리하는 규칙 기반 방식과 통계 방식을 혼용할 수도 있다. 하지만, 영어사전에 등재되어 있는 단어라 할지라도 4.1의 외래어 발음 변이에서 보이는 바와 같이 외래어이기 때문에 발생하는 발음변이가 있으므로, 사전의 발음열만을 채택할 경우 부작용이 있을 것으로 추측된다. 따라서 추후 연구에서는, 사전에 등재된 단어의 경우 원어의 발음을 채택하더라도 한국어 음소 체계에 맞게 변환하고, 이와 별개로 외래어 G2P로 발음을 생성하여 다중 발음열을 취하는 방식을 고려하면 보다 나은 성능을 기대할 수 있을 것으로 생각된다.

5. 결론

본 연구에서는 실제 어플리케이션 응용 환경에서 사용되는 외래어 어휘를 대상으로 한국인 화자의 외래어 발음 변이 양상을 분석하고, 이를 바탕으로 외래어 발음열을 자동으로 생성함으로써 자소-음소 변환 성능을 살펴보았다. 분석 결과, 먼저 한국인 화자의 외래어 발음 변이 양상으로 치경마찰음 ㅅ/s/의 경음화와 원순 모음 기/qi/와 게/we/의 원순성 약화를 관찰할 수 있었다. 또한 선/후행 음소 환경을 함께 살펴봄으로써 본 연구의 외래어 주요 발음 변이가 특정 음소 문맥 환경에서 실현되고 있음을 확인할 수 있었다. 이와 같은 특성을 반영하기 위해 문맥 환경 정보를 이용한 외래어 G2P를 이용하여, 음절 기반의 분절 정보를 이용한 외래어 자소-음소 방법을 제안하였다. 실험 결과, 음절 기반의 분절 정보를 이용한 제안 방법이 베이스라인보다 우수한 성능을 보였다. 특히, 앞뒤로 3개 음소씩 총 7-gram 환경에서 가장 우수한 성능을 보였다. 또한 외래어의 어휘 영역이 외래어 자소-음소 변환 성능에 영향을 미치고 있음을 밝혔다.

본 연구는 1) 고유명사를 포함하는 대규모 외래어 어휘를 대상으로 주변 문맥음소까지 고려하여 한국인의 외래어 발음 변이를 양적으로 분석하고, 2) 외래어에 대한 음성/음운론적 분석 결과를 반영함으로써 한국어에서의 외래어 G2P 성능을 향상시켰다는 점에서 의의를 갖는다. 하지만 본 연구에서는 자소-음소 변환 결과가 실제 음성 인식에서 어떤 영향을 미치는지는 보이지 못 하였다는 한계가 있다. 따라서 추후 연구에서는 본 연구에서 제시한 자소-음소 변환 방법이 음성 인식 성능에 미치는 영향에 대한 연구가 추가로 필요하다. 또한 중국어나 일본어와 같은 한자 기반의 외래어를 대상으로 하지 않았고, 3.5.2에서 언급한 바와 같이 자소-음소 변환의 입력으로 로마자만이 가능하여 한글로 표기된 외래어의 발음열 자동 생성을 다루지 않았다는 한계가 있으므로 이에 대한 추가 연구가 필요하다.

한편, 컴퓨터를 이용한 한국인 학습자의 영어 발음 교육과 같은 환경에서 자동 발음 오류 검출 및 피드백을 위해서는 오류 발음을 자동으로 예측할 수 있어야 한다(Bang et al., 2014). 특히 초급 수준의 학습자는 영어 단어를 발화할 때 외국어가 아닌 외래어 발음을 보일 공산이 있다. 따라서 한국인의 외래어 자소-음소 변환 결과가 학습자의 오류 발음열에 포함된다면, 자동 발음 오류 검출에 긍정적인 효과가 있을 것으로 기대할 수 있다.

감사의 글

이 논문은 2015년도 SKT의 ‘외래어 발음열 변환연구’ 과제의 일환으로 수행되었습니다.

참고문헌

- Bang, J., Lee, J., Lee, G. G., & Chung, M. (2014). Pronunciation variants prediction method to detect mispronunciations by Korean learners of English. *ACM Transactions on Asian Language Information Processing (TALIP)*. Vol. 13, No. 4, 1-21.
- Bisani, M. & Ney, H. (2008). Joint-sequence models for grapheme-to-phoneme conversion. *Speech Communication*. Vol. 50 No. 5, 434-451.
- Cha, J. E. (2007). A study on the standard pronunciation of the words of foreign origin and related matters. *Korean Linguistics*. Vol. 35, 363-390.
- (차재은 (2007). 외래어 표준 발음 문제에 대한 고찰. *한국어학* 35권, 363-390.)
- Choi, H. (2001). A research on loanword pronunciations. Seoul: The National Institute of the Korean Language.
- (최혜원 (2001). 외래어 발음 실태 조사. 국립국어연구원.)
- Chung, K. (2003). Transcription of Loanwords and Pronunciation of Foreign Languages. *Studies in foreign language education*. No. 17, 185-214.
- (정국 (2003). 외래어 표기법과 외국어 발음. *외국어교육연구* 17호, 185-214.)
- Jouvet, D., Fohr, D. & Illina, I. (2012). Evaluating grapheme-to-phoneme converters in automatic speech recognition context. *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 4821-4824.
- Kang, H. (1996). English loanword in Korean. *Studies in Phonetics, Phonology, and Morphology*. Vol. 2, 21-47.
- Kim, B., Lee, G. G. & Lee, J-H. (2002). Morpheme-based grapheme to phoneme conversion using phonetic patterns and morphophonemic connectivity information, *ACM Transactions on Asian Language Information Processing*. Vol. 1, No. 1, 65-82.
- Lee, E. K. (1998). Usage of Loanwords in Multimedia. *Saegugeosaenghwal*. Vol. 8, 41-59.
- (이은경 (1998). 방송 분야의 외래어 사용. *새국어생활* 8권, 41-59.)
- Lee, J. & Lee, G. G. (2009). A data-driven grapheme-to-phoneme conversion method using dynamic contextual converting rules for Korean TTS systems. *Computer Speech and Language*. Vol. 23, 423-434.
- Lee, K-N. & Chung, M. (2007). Morpheme-based modeling of pronunciation variation for large vocabulary continuous speech

- recognition in Korean. *IEICE Transactions on Information and Systems*. Vol. E90-D, No. 7, 1063-1072.
- Modipa, T. & Davel, M. H. (2010). Pronunciation modelling of foreign words for Sepedi ASR. *Proceedings of 21st Annual Symposium of the Pattern Recognition Association of South Africa (PRASA)*. 185-89.
- Ryu, H. & Chung, M. (2014). Pronunciation variations of loanwords produced by Korean speakers. *Proceedings of Conference of Korean Society of Speech Sciences (KSSS)*. 173-174.
(류혁수, 정민화. (2014). 한국인 화자의 외래어 발음 변이 양상. 한국음성학회 가을 학술대회 발표논문집, 141-142.)
- Schlippe, T., Ochs, S. & Schultz, T. (2012). Grapheme-to-phoneme model generation for Indo-European languages. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 4801-4804.
- Schlippe, T., Quaschnigk, W. & Schultz, T. (2014). Combining grapheme-to-phoneme enhanced pronunciation generation in low-resource scenarios. *Proceedings of 4th Workshop on Spoken Language Technologies for Under-resourced Languages (SLTU '14)*. 139-145.
- The National Institute of the Korean Language. (2015). The regulation concerning Hangeul transcription of loanwords. http://www.korean.go.kr/front/page/pageView.do?page_id=P000104&mn_id=97
(국립국어원. (2015). 외래어 표기법 규정. http://www.korean.go.kr/front/page/pageView.do?page_id=P000104&mn_id=97.)
- Wang, L. & Tong, R. (2014). Pronunciation modeling of foreign words for mandarin ASR by considering the effect of language transfer. *Proceedings of INTERSPEECH 2014*. 1443-1447.
- Yoon, K. & Brew, C. (2006). A linguistically motivated approach to grapheme-to-phoneme conversion for Korean. *Computer Speech and Language*. Vol. 20, 357-381.
- Yurn, G. D. (2006). In Justification of "Jjajangmyeon". *Korean Linguistics*. Vol. 30, 181-205.
(연규동 (2006). “짜장면”을 위한 변명. *한국어학* 30권, 181-205.)
- **류혁수 (Ryu, Hyuksu)**
서울대학교 언어학과
서울시 관악구 관악로 1
Tel: 02-880-9039
Email: oster01@snu.ac.kr
관심분야: 음성인식, 음성학, 컴퓨터 기반 언어교육
현재 서울대학교 언어학과 대학원 박사과정
 - **나민수 (Na, Minsu)**
서울대학교 인지과학협동과정
서울시 관악구 관악로 1
Tel: 02-880-9127
Email: dix39@snu.ac.kr
관심분야: 음성인식, 자연언어처리
현재 서울대학교 인지과학협동과정 대학원 박사과정
 - **정민화 (Chung, Minhwa)** 교신저자
서울대학교 언어학과
서울시 관악구 관악로 1
Tel: 02-880-9195 Fax: 02-882-2451
Email: mchung@snu.ac.kr
관심분야: 음성인식, 음성언어처리, 컴퓨터 기반 언어교육
현재 서울대학교 언어학과 교수