

영·한 음차 표기 성능 향상을 위한 음철법 기반 규칙 구축

Phonics-based Rules for Improving Performance of English-to-Korean Transliteration

김 민 정¹⁾ · 홍 금 원²⁾ · 박 소 영³⁾ · 임 해 창⁴⁾

Kim, Min-Jeong · Hong, Gumwon · Park, So-Young · Rim, Hae-Chang

ABSTRACT

This paper presents a method for constructing and using transliteration rules which are based on Phonics, an instructional method for speaking and writing English letters. Conventional approaches to automatic transliteration often focused on statistical methods. However, the construction or the collection of correct transliteration examples is always the bottleneck of the statistical transliteration model. Also, in practical domains where the collection of such data is very difficult, such as education and tourism, it is reasonable to build a system without much qualified data. Furthermore, compared with Korean orthography of borrowed foreign words, the proposed approach is much easier to construct, and can generate more refined rules. The experimentation result shows that the proposed approach can improve the performance of a statistical-based transliteration system.

Keywords: phonics, rule-based transliteration system, transliteration

1. 서 론

외국과의 교류가 빈번해지면서 신문 기사를 비롯한 한국어 문서에서 많은 외국어 단어들을 발음에 따라 한글로 표기하고 있다. 이렇게 원어의 소리에 따라서 문자나 말을 바꿔 쓰는 것을 음차(Transliteration)라고 한다[1]. 오래전부터 한국어는 한자 어들을 음차 표기하여 사용하여 왔으며, 최근에는 중국어, 일본어, 한국어에서 영어, 프랑스어, 독일어 등의 단어들을 음차 표기하고 있다[2]. 텔레비전, 디지털 카메라와 같은 일반 명사들과 밴쿠버, 샌프란시스코, 헬렌 켈러와 같은 인명과 지명을 포함하는 고유명사들도 음차 표기하여 쓴다.

음차 표기 결과물은 다양한 용도로 사용할 수 있다. 일단, 자

연어 처리 분야에서 바로 사용할 수 있다. 가장 대표적인 예가 교차 언어 검색으로, 질의어를 음차 표기해서 바로 다른 언어의 검색을 시도해볼 수 있다. 또한 외래어 그대로 많이 사용하는 의학, 약학 등의 전문용어 사전 구축도 가능하다. 이 이외에도 기계 번역 등의 분야에서 사용이 가능하다. 관점을 넓혀서 보면 음차 표기는 외국어 교육이나 관광, 문화 등의 분야에서도 사용이 가능하다.

실제로 영어와 같은 외국어의 교육에서 가장 처음에 하는 것은 글자 읽는 법을 습득하는 것이며 국내 영어 교육의 경우 음 철법(Phonics)을 도입하여 활용하고 있다. 음철법은 영어권 국가에서 5-6세 정도의 말을 알아듣고 이해하지만 글자를 읽고 쓰지 못하는 아이들에게 문자를 가르치기 위해 만든 교수법이다 [3]. 우리나라에서는 음철법을 초등학생 영어 교육에 도입하여 한글을 읽을 수 있지만 영어를 읽지 못하는 아이들의 교육에 사용하고 있다. 외래어의 사용이 급격히 증가하면서 문교부에서 외래어 표기법을 고시하였으며[4] 국립 국어원 홈페이지를 통해 외래어 표기 용례와 함께 공개하였다. 외래어 표기법은 영어뿐만 아니라 독일어, 일본어 등 21개국 언어의 표기 세칙까지 포함하고 있다. 음철법이나 외래어 표기법의 경우 모든 예외 사항을 다루고 있지는 않다. 하지만 각 글자나 발음기호 별로 가장 보편적인 발음을 다루고 있다고 할 수 있다. 따라서 음철

1) 고려대학교 mjkim@nlp.korea.ac.kr

2) 고려대학교 gwhong@nlp.korea.ac.kr

3) 상명대학교 ssoya@smu.ac.kr

4) 고려대학교 rim@nlp.korea.ac.kr, 교신저자

이 연구에 참여한 연구자의 일부는 '2단계 BK21사업'의 지원비를 받았음.

접수일자: 2009년 11월 9일

수정일자: 2009년 12월 10일

게재결정: 2009년 12월 10일

법이나 외래어 표기법을 기본으로 하고 예외가 되는 발음을 추가한다면 어느 정도 표준화된 음차 표기가 가능하다.

자동 음차 표기 문제를 접근하는 방법은 크게 사전 기반, 통계 기반, 규칙 기반 방법으로 나누어 볼 수 있다. 사전 기반 음차 표기 방법은 원어와 음차 표기한 단어의 쌍이 있는 사전에서 가장 비슷한 원어 단어를 찾고 그에 대응되는 음차 표기한 단어를 이용하는 방법이다[5]. 음차 표기를 수행할 수 있었을 때의 결과는 매우 정확한 편이며 자연스러운 음차 표기가 가능하나, 실제로 실생활에서 사용하는 모든 단어를 사전으로 구축하는 것은 불가능하며 이에 따라 미등록어 문제가 발생하게 된다는 단점이 있다. 따라서 많은 연구에서 통계 기반 방법으로 접근하였다.

통계 기반 음차 표기 방법은 외래어와 음차 표기한 한국어 쌍으로 구성된 말뭉치를 이용하는 방법으로, 최대 엔트로피 모델을 이용한 방법[1], 결정 트리를 이용한 방법[6], 메모리 기반 학습과 결정 트리를 사용한 방법[7], 음운 패턴을 이용한 방법[8], 음소 단위 n-gram 음차 표기 확률을 이용한 방법[2,9] 등이 있다. 통계 기반 음차 표기 모형은 사전 기반 음차 표기 모형에 비해 자료 부족 문제에 매우 강건한 장점을 가지고 있다. 하지만 단순히 통계 자료만으로 음차 표기를 하다보면 학습에 사용한 데이터와 실제 음차 표기를 수행해야 하는 단어의 성격이다를 경우 전혀 엉뚱한 발음도 나올 수 있다. 예를 들어, 인명, 지명 표기 용례로 학습한 통계 기반 음차 표기 모형으로 일반 단어의 음차 표기를 수행했을 때 ‘Sunday’를 ‘순데이’로, ‘Sale’을 ‘살레’로 표기하였다. 자연어 처리 분야에서 음차 표기 결과물을 사용할 때에는 어느 정도의 오류가 있어도 사용이 가능했다. 하지만 외국어 교육이나 문화, 관광과 같은 분야에 사용하고자 하면 어느 정도의 비용을 감수하고 사람이 개입하여 음차 표기 성능을 개선할 수 있어야 한다. 통계 기반 방법의 성능을 향상시키려면 많은 음차 표기의 예를 포함하는 좋은 품질의 말뭉치를 다량으로 수집 혹은 구축하는 것이 필요하며, 이것은 시간과 비용이 많이 듈다. 또한 상대적으로 용례 수집이 어려운 비영어권 단어나 신조어의 처리가 어려우며, 좋은 품질의 음차 표기 자료 없이 음차 표기 시스템의 품질을 향상시키는 것이 힘들다.

규칙 기반의 음차 표기 방법은 사람이 수동으로 구축한 규칙이나 말뭉치에서 자동으로 추출한 규칙을 이용하여 음차 표기를 수행한다. 규칙을 구축하는데 시간과 비용이 많이 들기 때문에 한국어 음차 표기를 위한 이전 연구는 통계 기반 방법을 위주로 이루어졌으나, 용례를 많이 수집하기 어려운 경우나 신조어 등의 처리를 위해서는 직관적으로 관리할 수 있는 규칙 기반 방법이 적합하다. 마침 전문가들에 의해 만들어진 음절법과 외래어 표기법이 있고, 이를 이용하면 초기 음차 표기 규칙을 구축해볼 수 있다.

기존 연구에서 규칙 기반 음차 표기를 할 때에는 주로 외래

어 표기법을 사용하였다[10]. 외래어 표기법을 적용하기 위해서는 외래어를 발음으로 전환하고 이를 다시 한국어로 변환하는 과정이 필요하다. 이렇게 두 번 변환을 하게 되면 첫 번째 단계의 오류가 두 번째 단계에도 전파가 되어 전체적인 성능을 하락시키게 된다. 음절법은 외래어의 철자를 고려하여 직접 한국어로 변환하도록 되어 있으므로 이러한 오류 전파의 가능성성이 없다. 또한 음절법은 학습 대상을 외국어를 배우려는 일반인으로 하고 있으므로 많이 사용하는 단어에 대해 외래어 표기법보다 상세하게 발음을 다루고 있고, 단어 내의 모음 개수나 주변 음소까지도 고려하여 음차 표기를 하도록 되어 있다. 그러므로 음절법을 적용할 경우 외래어 표기법보다 자세한 경우를 다루는 규칙을 작성할 수 있으며, 보다 정확한 결과를 얻을 수 있다. 따라서 본 연구에서는 음절법을 이용한 음차 표기를 시도한다.

본 논문에서는 기존의 음절법을 그대로 적용할 수 있도록 규칙의 형태를 정의하고 이 규칙들을 사용하여 음차 표기를 하는 규칙 기반 음차 표기 시스템을 구축하였다. 그리고 음절법을 적용하여 규칙을 작성하고, 이를 적용한 결과를 분석하였다. 또한 통계 기반 방법의 견고함과 규칙 기반 방법의 정교함을 결합해 보기 위해 두 가지 모형을 결합하는 방법도 시도한다.

본 논문은 다음과 같이 구성된다. 2장에서는 음절법을 이용한 음차 표기 규칙에 대해 설명하며, 3장에서는 이러한 규칙을 이용하여 음차 표기를 수행하는 규칙 기반 음차 표기 시스템에 대해 기술한다. 4장에서는 이 규칙과 시스템을 이용한 음차 표기 결과를 분석하며, 마지막으로 5장에서는 결론을 맺는다.

2. 음절법을 활용한 음차 표기 규칙

음차 표기는 외국어 단어의 철자로부터 발음을 추측하여 한글로 표기하는 방법인 눈말 표기 방법과 외국어 단어의 발음기호에 근거하여 한글로 표기하는 입말 표기 방법이 있다[11]. 음차 표기 시스템의 입력은 외국어 단어이므로, 눈말 표기 방법은 외국어 단어를 직접 음차 표기하게 되나 입말 표기 방법은 외국어 단어를 발음기호로 변환한 후에 음차 표기를 하게 된다[9]. 직관적으로 사람이 외국어 단어를 읽는 방법은 입말 표기 방법에 더 가까우나, 음차 표기 시스템에 이를 구현할 때에는 입말 표기 방법의 경우 철자를 발음 기호로 변환하는 단계에서 발생하는 오류가 발음 기호를 사용하여 음차 표기하는 단계에 영향을 미치기 때문에 성능이 더 떨어지게 된다[10,12]. 기존 연구에서 규칙 기반 음차 표기를 할 때 주로 사용한 것은 외래어 표기법이었다[10]. 하지만 외래어 표기법은 발음 기호에 따라 읽는 것이므로 입말 표기 방법이며 음절법은 영어 철자를 보고 그대로 읽는 것이므로 눈말 표기 방법이다. 본 논문에서는 오류를 줄이기 위해 눈말 표기 방법을 따르며, 음절법을 이용하여 규칙을 작성하도록 하였다.

본 논문에서 정의하는 음차 표기 규칙의 형식은 <표 1>과 같다. 우선 음차 표기 규칙은 크게 규칙을 적용하기 위한 조건을 제시하는 조건부와 조건을 충족했을 때 필요한 연산을 다른 실행부로 나눌 수 있다. 음차 표기를 적용할 때 변환 대상인 음소와 함께 주변 음소를 고려하므로 조건부는 다시 크게 변환 대상 음소 정보와 주변 음소 정보로 나눌 수 있으며 실행부는 음차 표기 규칙 정보만으로 구성된다.

2.1 변환 대상 음소 정보

본 논문에서 구축하는 규칙 기반 음차 표기 시스템은 음차 표기를 하고자 하는 단어를 음소열로 분해하고 배열에 저장한 후, 배열에서 차례로 음소를 읽으면서 해당 음소에 맞는 규칙을 찾아서 적용한다. 이 때, 음차 표기하고자 하는 음소가 자음인가 혹은 모음인가의 여부에 따라 참고하게 되는 정보에 차이가 있기 때문에 규칙의 모호성을 줄이기 위해 자음/모음 여부를 규칙에 포함하였다. <표 1>에서 ‘문장 내 소리 나는 모음 개수’와 ‘문장 내 몇 번째 모음인지에 대한 정보’의 경우 모음을 처리할 때만 필요한 정보이다.

표 1. 음차 표기 규칙 형식

Table 1. The specification of transliteration rules

구분	분류	규칙에 포함하는 내용
조건부	변환	음소 (source)
	대상	자음/모음 여부 (jamo)
	음소	단어 내 소리 나는 모음 개수 (moc)
	정보	단어 내에서 몇 번째 모음인지에 대한 정보 (moi)
	주변	음소 (context)
실행부	주변	+ 주변 음소의 언어
	음소	+ 변환 대상 음소와의 상대적 위치
	정보	주변 음소의 자음/모음 여부 (contjamo)
		단어의 끝인지 여부 (end)
	음차 표기 규칙 정보	음차 표기 결과가 되는 한국어 음소 (target) + 음차 표기 결과가 되는 한국어 음소의 종류 + 음차 표기 후 한국어 음소의 위치 전녀별 한국어 음소 배열 색인 (koi) 전녀별 외국어 음소 배열 색인 (eni)

- source:** 음차 표기하고자 하는 음소. 영어 단어의 경우 알파벳 한 글자가 된다.
- jamo:** 음차 표기하고자 하는 음소가 자음과 모음 중 어느 것인지에 관한 정보. 자음일 경우 ja로 표기하며 모음일 경우 mo로 표기한다.
- moc:** 단어 내에 소리 나는 모음이 몇 개인지를 나타내는 정보. 실제로 필요한 정보는 단어 내에 소리 나는 모음이 1개 인지와 2개 이상인지를의 이진분류이므로, 1개일 경우는 1로 표기하고 2개 이상일 경우는 2로 표기한다. 단어 내의 모음 개수가 상관없는 경우나 자음 규칙에는 0으로 표기한다.
- moi:** 현재 음차 표기하고자 하는 음소가 단어 내에 몇 번째

로 소리 나는 모음 음소인지에 관한 정보. moc와 마찬가지로 실제로 필요한 정보는 첫 번째인가 아닌가의 여부이므로 첫 번째 소리나는 모음에 관한 조건에서는 1로 표기하고, 두 번째나 그 이후의 모음일 경우에는 2로 표기한다. 몇 번째 모음인지 상관없는 경우나 자음 규칙의 경우에는 0으로 표기한다.

2.2 주변 음소 정보

음차 표기를 위해 주변 음소가 무엇인지를 고려해야 할 때가 있다. 예를 들어, ‘knight’라는 영어 단어는 ‘나이트’라고 읽게 된다. 이는 맨 앞의 ‘k’가 ‘n’과 붙어 나올 때에는 묵음이 되기 때문이다. 또한 영어에서 대부분의 자음이 단어의 맨 뒤나 다른 자음의 앞에 올 때에는 중성 ‘-’를 삽입하는 등 주변 음소에 따라 발음이 바뀌거나 생략되거나 추가되는 경우가 생기게 된다. 이런 정보를 규칙으로 표현할 수 있게 하기 위해 조건부에 주변 음소 정보를 추가하였다.

- context:** 음차 표기하고자 하는 음소의 주변에 고려해야 하는 다른 음소 정보. 영어 단어의 경우 알파벳 한 글자가 된다. 단, 앞이나 뒤에 오는 음소가 단지 자음과 모음 중 어느 것에 해당하는지의 정보만 필요한 경우에는 0으로 표기한다. 또한 ‘-ing’와 같이 몇 개 음소 뒤에 단어의 끝이 오는 경우에는 ‘n’을 음소 위치에 표기한다.
- context의 언어 정보:** 대부분의 경우 외국어의 문맥을 보게 되나 영어의 알파벳 ‘l’을 음차 표기하면서 앞 음절이 모음으로 끝날 경우 받침으로 ‘ㄹ’을 삽입해야 하므로 한국어 배열을 조건으로 사용할 경우도 있다. 따라서 context항에 언어를 추가한다. 고려해야 하는 음소가 외국어일 경우에는 en으로 표기하고 고려해야 하는 음소가 한국어일 경우에는 ko로 표기한다.
- context의 위치 정보:** 현재 음차 표기하고자 하는 음소 (source항의 음소)보다 앞에 나오는 음소일 경우에는 음수로 표기하고, 현재 음차 표기하고자 하는 음소보다 뒤에 나오는 음소일 경우에는 양수로 표기한다.
- contjamo:** 음차 표기하고자 하는 음소의 주변 음소, 특히 context로 지정한 문맥의 바로 앞, 뒤 음소가 자음과 모음 중 어느 것인가에 관한 정보. 예를 들어, 영어 단어의 경우 ‘s’ 뒤에 자음이 올 경우에는 ‘스’라고 음차 표기하게 된다. 이렇게 자음의 경우 뒤에 다른 자음이 오면 ‘-’를 삽입하게 되는 경우가 있으므로 주변 음소의 자음/모음 여부를 필요로 하게 된다. 주변 음소가 자음일 경우에 적용해야 하는 규칙에는 ja라고 표기하며, 주변 음소가 모음일 경우에 적용해야 하는 규칙에는 mo라고 표기한다. 이러한 정보가 상관없을 경우에는 0으로 표기한다. 또한 보고자 하는 음소가 앞의 음소일 경우 -ja, 뒤의 음소일 경우 +ja와 같이 표시한다. 또한 만약 보고자 하는 음소의 이전과 이후를 모두 고려하고자 할 경우에는 -mo,+mo와 같이 병렬로 표기한다.
- end:** 음차 표기하고자 하는 음소가 단어의 마지막에 나오는 음소인지의 여부. 예를 들어, 영어 단어의 경우 ‘s’ 가 단어의 마지막 음소일 경우 ‘스’라고 음차 표기하게 된다. 이렇

게 마지막 음소가 자음일 경우 ‘-’를 삽입하게 되는 경우를 위해 마지막 음소인지 여부를 표기한다. 마지막 음소일 때 적용하는 규칙에는 yes라고 표기하며 마지막 음소가 아닐 때 적용하는 규칙에는 no라고 표기한다. 이러한 정보가 상관없을 경우에는 0으로 표기한다.

2.3 음차 표기 규칙 정보

2.1절과 2.2절에서 설명한 조건부가 모두 만족이 되었을 때 수행하게 되는 음차 표기 규칙을 포함하는 부분이다.

- target:** 음차 표기 결과가 되는 한국어 음소. 자음과 모음 모두 포함한다.
- target의 종류 정보:** 음차 표기 결과가 되는 한국어 음소의 초성, 중성, 종성 정보. 자음의 경우 초성과 종성을 다르게 입력해야 하므로 추가로 표기하도록 한다. 초성은 1, 중성은 2, 종성은 3으로 표기한다.
- target의 위치 정보:** 음차 표기 결과 한국어 음소의 위치 정보. 한 번에 여러 개의 한국어 음소를 생성해야 할 경우가 있으므로 위치 정보를 추가로 표기한다. 2.2절의 context의 위치 정보와 같이 현재 입력하는 한국어 음소의 위치를 기준으로 앞의 음소를 바꾸게 되면 음수 값으로 표기하며, 현재 입력하는 음소보다 뒤에 음소를 추가하게 되면 양수 값으로 표기한다.
- koi:** 다음 연산에서 건너뛸 한국어 음소 배열의 위치. 여러 한국어 음소를 추가했을 경우 다음 연산에서 사용할 배열의 위치를 지정해주어야 한다. 한 개의 음소를 처리했을 때 배열의 위치를 한 칸씩 이동하는 것이 기본이므로 한 개 이상 이동해야 할 경우에 추가로 이동해야 하는 범위를 양수로 표기한다. 추가로 이동할 필요가 없을 경우에는 0으로 표기 한다.
- eni:** 다음 연산에서 건너뛸 외국어 음소 배열의 위치. 예를 들어, 영어 단어 ‘church’에서의 ‘ch’, ‘wood’에서의 ‘woo’와 같이 한 번에 처리해야 할 음소가 많을 경우에는 규칙 조건부의 주변 음소 정보에서 포함하고 다음 연산에서 건너뛰게 된다. 한 개의 음소를 처리했을 때 배열의 위치를 한 칸씩 이동하는 것이 기본이므로 koi와 같이 한 개 이상 이동해야 할 경우에 추가로 이동해야 하는 범위를 양수로 표기한다. 마찬가지로 추가로 이동할 필요가 없을 경우에는 0으로 표기한다.

2.4 음철법을 이용한 규칙 구축

<표 2>는 실제로 음철법을 적용하여 영어를 한국어로 음차 표기하도록 만든 규칙의 예이다. 음철법은 영어 교육과 관련하여 학습지나 학원에서 많이 사용하나, 외부로 공개하지는 않으므로 본 논문에서는 음철법에 대해 상세히 다루고 있는 커뮤니티⁵⁾와 지식 질의응답 사이트⁶⁾를 참고하여 규칙⁷⁾을 작성하였

다. 가장 긴 문맥을 본 경우는 ‘young’를 음차 표기하는 것으로, ‘you’는 ‘유’로 발음되고 ‘young’는 ‘영’으로 발음해야 하므로 ‘y’를 기준으로 뒤로 4개 음소까지 고려하도록 작성하였다.

<표 2>에서 ‘모음 규칙’은 영어 단어에서 현재 음차 표기하고자 하는 음소가 ‘e’이고, 단어 내에서 소리 나는 모음이 2개 이상인데 그중 첫 번째 모음이 아닐 경우이다. 또한 바로 뒤의 음소가 ‘a’일 경우 ‘ea’를 ‘ㅏ’로 음차 표기하는 규칙이다.

<표 2>의 ‘자음 규칙’은 현재 음차 표기하고자 하는 음소가 ‘d’이고 단어의 마지막 음소일 경우 ‘드’로 음차 표기하는 규칙이다.

표 2. 음차 표기 규칙의 예

Table 2. Examples of transliteration rules

규칙의 종류	음철법을 적용한 실제 규칙의 예
모음 규칙	source=e jamo=mo moc=2 moi=2 context=a(en,+1) contjamo=0 end=0 target=ㅣ(2,0) koi=0 eni=+1
자음 규칙	source=d jamo=ja moc=0 moi=0 context=0 contjamo=+ja end=yes target=ㄷ(1,0),-(2,+1) koi=+1 eni=0

이렇게 규칙을 작성할 때에는 규칙을 적용할 때 모호성을 줄이기 위해 다음과 같은 두 가지 원칙을 두었다.

- 조건부의 모든 조건이 만족할 경우에만 실행부의 연산을 수행한다. 즉, AND 연산으로 가정하며 OR 연산이 될 경우에는 조건부를 다르게 하여 각기 다른 규칙으로 구축한다.
- 규칙의 적용 순서는 규칙 파일 내에 상위에 있는 규칙부터 순서대로 적용하며, 어느 한 개의 규칙의 조건부가 만족될 경우 실행부의 연산을 수행하고 다음 음소로 넘어간다. 따라서 규칙 파일 내의 규칙의 순서가 중요하게 되며 가장 구체적인 조건을 포함하는 규칙을 상위에 놓는 것이 중요하다.

3. 규칙 기반 음차 표기 시스템

2장에서 정의한 규칙을 적용해보기 위해 본 연구에서는 <그림 1>과 같은 시스템을 구축한다. 이 시스템은 입력 단어를 음소 단위로 분해하고 결과 음소 열을 음절 단위로 합성하는 모듈, 규칙을 적용하기 위해 필요한 정보로 단어 내 소리 나는 모음 개수를 세는 모듈, 자음 규칙을 적용하는 모듈, 모음 규칙을 적용하는 모듈, 모음으로 시작하는 음절에 초성을 삽입하는 모듈로 구성되어 있다.

5) <http://cafe.naver.com/www05cokr>

6) http://kin.naver.com/detail/detail.php?dir_id=111401&docid=4080130

7) 자음 규칙보다는 모음 규칙이 복잡하며 수가 많다. 본 논문

에서 사용한 규칙은 150개이며, 실험 후에 r등의 자음에서의 몇 가지 예외사항을 추가하고 상세화한 규칙을 추가하여 총 226개의 규칙을 부록에서 표의 형태로 공개한다.

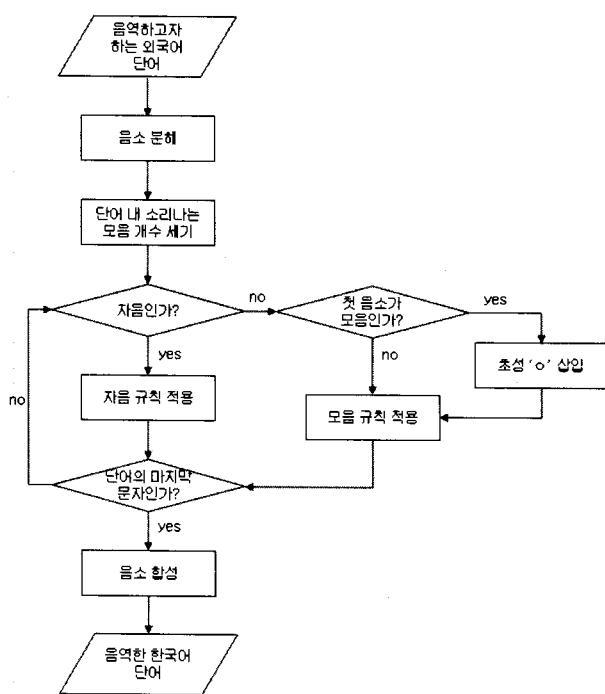


그림 1. 규칙 기반 음차 표기 시스템 구성도

Figure 1. The rule-based transliteration system architecture

3.1 음소 분해 및 음소 합성

한국어에서는 여러 개의 음소가 음절을 구성하며, 음차 표기는 각 음소별로 이루어진다. 음소 분해 모듈은 외국어 단어가 입력되면 이를 각 글자로 분해하여 입력 배열에 집어넣는다. 음소 합성 모듈은 출력 배열에 있는 한국어 음소열을 음절 단위로 합성한다.

3.2 단어 내 소리 나는 모음 세기

모음 규칙에서는 단어 내에 소리 나는 모음이 몇 개이며, 그 중 현재 모음이 몇 번째 모음인지에 따라 다르게 음차 표기가 되는 경우가 있다. 단어 내 소리 나는 모음을 세기 위해 우선 각 음소를 자음과 모음을 구분한 후에 “ea”, “woo”와 같이 연이어 나오는 모음을 한 개의 모음으로 세도록 하였다.

3.3 자음 음차 표기 규칙 적용

자음 규칙 적용 모듈에서는 단어 내의 모음 개수나 현재 음소가 단어 내에서 몇 번째 모음인지에 대한 조건을 확인하지 않는다. 대신 뒤에 자음이 오는 경우가 있는지, 혹은 현재 음차 표기하고자 하는 음소가 단어의 마지막 음소가 아닌지를 반드시 확인한다.

3.4 모음 음차 표기 규칙 적용

모음 규칙 적용 모듈에서는 단어 내의 모음 개수와 현재 음소가 단어 내에서 몇 번째 모음인지에 대한 조건을 확인하는

경우가 많으며, 음절 내에서 첫 음소가 모음인지도 확인하게 된다.

3.5 음차 표기 규칙을 적용하여 음차 표기를 하는 예
 입력된 단어가 “Mead”라고 가정하자. ‘m’의 경우 ‘mn’이나 ‘mm’의 경우에만 예외를 처리하는 규칙이 있고, 그 이외에는 앞에 모음이 있을 때에는 종성 ‘ㅁ’이 되며 뒤에 모음이 올 경우에 초성 ‘ㅁ’으로 발음한다. 따라서 순차적으로 ‘m’은 초성 ‘ㅁ’이 되고, 앞의 <표 2>의 규칙에 따라 ‘ea’는 ‘ㅣ’이 되며, 단어의 마지막 음소인 ‘d’는 ‘ㄷ—’가 된다. 이렇게 차례로 변환된 음소를 마지막에서 합성을 하게 되면 음차 표기 결과로 “미드”를 출력하게 된다.

4. 실험 결과 및 분석

4.1 평가 집합

본 논문에서는 국립 국어원에 공개되어 있는 외래어 표기 용례 자료 중 로마자 인명과 지명을 평가 집합으로 사용하였다. 단, “Prince Andrew, 앤드루 왕자”와 같이 여러 단어로 되어 있는 경우 띄어쓰기를 기준으로 분리를 하고 “Prince”를 “왕자”로 표기한 것은 음차 표기가 아닌 번역이므로 이런 경우는 수동으로 찾아서 평가 집합에서 제거하였다. 최종 평가 집합으로 사용한 단어는 총 15,651개이다.

4.2 평가 척도

본 논문에서는 NEWS 2009 Machine Transliteration Shared Task⁸⁾에서 사용한 평가 척도[13] 중 평균 F-measure를 평가 척도로 사용하며, 한글은 초성, 중성, 종성을 모아쓰기를 하므로 이를 분해하여 음소 단위로 F-measure를 측정한다.

표 3. 규칙 기반 음차 표기 결과 F-measure별 분포
 Table 3. The distribution of rule-based transliteration result

F-measure 범위	개수	비율
0 (전부 틀린 경우)	12	0.08%
0.1 이상 0.2 미만	22	0.14%
0.2 이상 0.3 미만	123	0.79%
0.3 이상 0.4 미만	259	1.65%
0.4 이상 0.5 미만	583	3.73%
0.5 이상 0.6 미만	1514	9.67%
0.6 이상 0.7 미만	2491	15.92%
0.7 이상 0.8 미만	2858	18.26%
0.8 이상 0.9 미만	4059	25.93%
0.9 이상 1.0 미만	711	4.54%
1 (정확히 맞은 경우)	3019	19.29%
총 합계	15651	

8) <http://www.acl-ijcnlp-2009.org/workshops/NEWS2009/pages/sharedtask.html>

F-measure를 계산하기 위해서는 최장 공통 음소열의 길이를 구하고, 이를 이용하여 정확률과 재현율을 계산한다. 최장 공통 음소열의 길이는 정답 단어와 시스템 결과 각각의 단어 길이와 두 단어 간의 편집 거리를 이용하여 구한다. 정답 단어를 r , 시스템의 결과를 c 라고 하고 편집 거리를 $ED(c, r)$ 이라고 했을 때 최장 공통 음소열을 구하는 수식은 다음 수식 (1)과 같다.

$$LCS(c, r) = 1/2(\text{length}(c) + \text{length}(r) - ED(c, r)) \quad (1)$$

최장 공통 음소열을 이용하여 단어 i 의 정확률(P_i)과 재현율(R_i)을 측정하면 각각 다음 수식 (2), (3)과 같이 된다.

$$P_i = \frac{LCS(c_i, r_i)}{\text{length}(c_i)} \quad (2)$$

$$R_i = \frac{LCS(c_i, r_i)}{\text{length}(r_i)} \quad (3)$$

따라서 단어 i 의 음소 단위 F-measure는 다음 수식 (4)와 같이 측정할 수 있다.

$$F_i = 2 \frac{R_i \times P_i}{R_i + P_i} \quad (4)$$

음소 단위 F-measure의 값이 1에 가까울수록 시스템의 결과가 정답과 비슷함을 뜻하며, F-measure 값이 1이 되면 정확히 일치함을 뜻한다.

4.3 규칙 기반 음차 표기 모형 적용 실험 결과 및 분석
 본 논문에서 제안한 규칙 기반 모형으로 외래어 표기 용례의 인명 자료와 지명 자료 15,651개의 단어를 음차 표기해본 결과, F-measure 기준으로 0.766을 기록하였다. 이중 F-measure 0.5 미만은 999개였으며, 3,019개 단어는 정확히 음차 표기하였다. 또한 오류를 분석해본 결과 <표 4>의 예와 같이 영어 이외의 언어권의 고유 명사가 F-measure를 기준으로 하위권에 있음을 확인할 수 있었다. 이는 본 논문에서는 영어를 대상으로 한 음 철법만을 대상으로 규칙을 구축하였으나, 외래어 표기 용례 자료는 영어 이외의 로마자 표기 언어도 포함하고 있기 때문이다. 이런 단어의 경우 그에 해당하는 언어의 규칙을 추가하면 음차 표기가 가능할 것으로 판단된다.

표 4. 음차 표기 오류의 예
Table 4. Examples of transliteration errors

외국어	정답	시스템 결과	F-measure
Wit	빗	위트	0.00
Beauchamp	보샹	비우참프	0.14
Judaea	유대	주데이아	0.17
QianXuesen	첸쉐센	키안수이즌	0.30

4.4 음철법 기반 규칙의 효과

<표 5>는 본 논문에서 제안한 규칙 기반 음차 표기 모형을 이용해서 통계 기반 모형을 보완해보는 실험을 한 결과이다 [14]. 규칙 기반 음차 표기 모형은 규칙이 있는 경우에는 정확하지만 규칙이 작성되지 않은 음소의 경우에는 잘못 음차 표기할 수 있다. 통계 기반 모형은 규칙을 사람이 작성할 필요 없이 응용률을 직접 이용하여 학습하므로 응용에 나타난 음소에 대해서는 음차 표기가 가능하다.

통계 기반 모형은 음소 단위 정렬에는 GIZA++[15]를 사용하였고, 최적 후보를 찾기 위해서는 통계 기반 구 단위 기계번역 시스템인 MOSES[16]를 이용하였다. 구 기반 번역 시스템을 사용했기 때문에 본 연구에서는 음소열의 최대 길이를 3으로 설정하였다. 또한 통계 기반 모형은 번역 모형과 언어 모형에 따라 여러 후보가 가능하다. 따라서 규칙 기반 모형의 결과를 참고하여 통계 기반 모형의 결과를 재순위화 하였다. 만약 규칙 기반 모형의 결과와 같은 결과가 통계 기반 모형의 결과에 있을 경우에는 점수에 임의의 값을 더해주었으며, 규칙 적용 통계 기반 모형의 결과에 없을 경우에는 특정 순위에 규칙 기반 모형의 결과를 삽입하였다. 이때 더해주는 가중치와 삽입하는 순위는 실험에 의해 결정하였다.

<표 5>는 4.1절의 외래어 표기 용례 자료를 이용하여 통계 기반 모형을 학습하고 NEWS2009에서 실험 집합으로 사용한 외래어 인명, 지명 표기 용례 989개를 대상으로 실험한 결과이다. 학습에 사용한 용례 자료와 실험에 사용한 용례 자료 모두 인명과 지명에 관한 자료였으므로 성격이 비슷하며, 통계 기반 모형의 성능이 매우 높았다. 영어 음차 표기 규칙만 사용하여 영어 이외의 외래어로 인한 오류가 많았음에도 불구하고 규칙을 사용하였을 때 성능이 올라감을 확인할 수 있었다. 이는 규칙이 통계 기반 모형의 오류를 수정할 수 있었음을 의미한다.

표 5. 통계 기반 음차 표기 방법에 음철법 기반 규칙을 결합했을 때의 실험 결과

Table 5. Experimental result of adding phonics-based rules to statistical-based transliteration systems

통계 기반 음차 표기 모형	평균 F-measure	성능 향상 폭
규칙 제외	0.868	-
규칙 포함	0.887	2.19%

5. 결 론

본 논문에서는 영어-한국어 음차 표기를 위해 음철법을 적용하는 방법을 제안하였다. 음철법은 영어권에서 영어를 읽고 쓰는 법을 가르치기 위해 개발된 교사법으로, 국내에서는 초등학생의 영어 교육을 위해 사용하고 있다. 특히 우리나라에서의 음 철법은 영어 단어를 보고 한국어로 발음을 표기할 수 있도록

가르치고 있으므로 음차 표기에 적용하기에 적합하다. 하지만 이전 연구에서는 외래어 표기법만을 적용하였다. 이에 본 연구에서는 음절법을 그대로 적용 가능한 규칙의 형태를 정의하고 규칙 기반 음차 표기 시스템을 구현하여 음절법을 그대로 음차 표기에 적용하여보았다. 적용해서 실험해 본 결과, 음절법만으로도 실패한 단어 없이 모든 외래어를 음차 표기할 수 있었다.

또한 본 논문에서 제안한 음절법 기반의 규칙 시스템을 구현위 번역 모델을 이용한 기존의 통계 기반 음차 표기 방법을 보완하는데 적용해본 결과, 통계 기반 방법에서 사용한 학습 말뭉치와 평가 집합의 성격이 매우 유사함에도 성능이 향상되었으며, 이를 통해 정확률과 재현율을 모두 개선할 수 있는 가능성을 보였다. 매우 간단한 방법을 통해 보완을 하였고, 이후에는 이런 보완 방법을 좀 더 다양하게 연구하여 좀 더 정교한 보완 방법을 찾아보아야 할 것이다.

본 논문에서 사용한 규칙은 영어-한국어의 음차 표기에 관한 규칙만 있어 다른 언어에 어원을 둔 단어의 경우 상대적으로 높은 오류를 보였다. 본 논문에서는 영어만을 대상으로 하였으나 중국어의 병음 표기, 프랑스어 등 다른 언어의 교육 과정을 반영하면 다른 언어-한국어 간의 음차 표기로도 확장이 가능하다. 다른 언어들에 대한 규칙으로 확장하여봄과 동시에 원어에 따라 다른 규칙을 적용하는 연구가 필요할 것이다. 또한, 현재에는 규칙에서 원어의 강세나 형태소 분석을 고려하지 않고 있는데 추후에는 이런 것도 고려할 수 있도록 규칙을 확장한다면 더욱 정확도를 높일 수 있을 것이다.

참 고 문 헌

- [1] Kim, T.I. (1999). "English-Korean transliteration model using maximum entropy model for cross language information retrieval", M.S. thesis, Sogang University.
(김태일, (1999). "최대 엔트로피 모델을 이용한 다국어 정보 검색에서의 영-한 음차 표기 모델", 서강대학교 석사학위 논문.)
- [2] Li, H., Zhang, M., Su, J. (2004). "A joint source-channel model for machine transliteration", *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, 2004, pp. 159-166.
- [3] Wikipedia, "Phonics", <http://en.wikipedia.org/wiki/Phonics>, (2009).
- [4] Ministry of Education & Human Resources Development, (1986). "Loan Words Notation", http://www.korean.go.kr/08_new/index.jsp.
(문교부, (1986). "외래어 표기법", http://www.korean.go.kr/08_new/index.jsp)
- [5] Lee, D.G. (2009). "Building English-to-Korean Transliteration Dictionary Based on Pronouncing Dictionary", *Phonetics and Speech Sciences*, Vol. 1, No. 3, pp. 105-110, September.
(이도길, (2009). "발음 사전에 기반한 영·한 음차 표기 사전의 구축," 말소리와 음성과학, 제 1권 제 3호, pp. 105-110.)
- [6] Kang, B.J., Choi, K.S. (2000). "Automatic transliteration and back-transliteration by decision tree learning", *Proceedings of the 2nd International Conference on Language Resources and Evaluation*.
- [7] Oh, J.H., Choi, K.S. (2005). "An English-to-Korean transliteration model based on grapheme and phoneme", *Journal of KISS : software and applications*, Vol. 32, No. 4, pp. 312-326, April.
(오종훈, 최기선, (2005). "자소 및 음소 정보를 이용한 영어-한국어 음차표기 모델", 정보과학회논문지:소프트웨어 및 응용, 제32권 제4호, pp.312-326.)
- [8] Kang, I.H., Kim, G.C. (1999). "English-to-Korean transliteration using multiple unbounded overlapping phonemes", *Proceedings of Annual Conference on Human and Language Technology*, pp. 50-54.
(강인호, 김길창, (1999). "복수 음운 정보를 이용한 영·한 음차 표기", 제11회 한글 및 한국어 정보처리 학술대회 논문집, pp. 50-54.)
- [9] Lee, J.S., Choi, K.S. (1997). "Automatic foreign word transliteration model for information retrieval", *Proceedings of the Korean Society for Information Management Conference*, pp.17-24.
(이재성, 최기선, (1997). "정보검색을 위한 외래어 자동표기 모델", 한국정보관리학회 제4회 학술대회 논문집, pp.17-24.)
- [10] Oh, J.H., Choi, K.S. (2002). "An English-Korean Transliteration Model Using Pronunciation and Contextual Rules", *Proceedings of the 19th International Conference on Computational linguistics*, Vol. 1, pp.1-7.
- [11] Lee, H.B. (1979). "On transliterating loan words by the Korean alphabet - A critical appraisal of the revised Koreanization system", *Language Research*, Vol. 15, No. 1, pp. 39-59.
(이현복, (1979). "외래어 표기법 개정 시안의 문제점", 어학 연구, 제15권, 제1호, pp.39-59.)
- [12] Oh, J.H. (2005). "English term recognition and English-Korean translation knowledge acquisition for terminology management", Ph.D. thesis, KAIST.
(오종훈, (2005), "전문용어관리를 위한 영어 전문용어인식 및 영-한 번역지식의 획득", 한국과학기술원 박사학위 논문.)
- [13] Li, H., Kumaran, A., Zhang, M., Pervouchine V. (2009). "Whitepaper of NEWS 2009 machine transliteration shared task", *Proceedings of the 2009 Named Entities Workshop, ACL-IJCNLP 2009*, pp.19 - 26.
- [14] Hong, G., Kim, M.J., Lee, D.G., Rim, H.C. (2009). "A hybrid approach to English-Korean name transliteration", *Proceedings of the 2009 Named Entities Workshop, ACL-IJCNLP 2009*, pp. 108 - 111.
- [15] Och F.J., Ney, H. (2003). "A systematic comparison of various statistical alignment models", *Computational Linguistics*, Vol. 29, Issue 1, pp. 19-51.
- [16] Koehn, P., Hoang, H., Birch, A., Callison-Burch, C., Federico, M., Bertoldi, N., Cowan, B., Shen, W., Moran, C., Zens, R., Dyer, C., Bojar, O., Constantin, A., Herbst, E. (2007). "Moses: Open Source Toolkit for Statistical Machine Translation", *Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions*, 2007, pp. 177-180.

• 김민정 (Kim, Min-Jeong)

고려대학교 컴퓨터 · 전파통신공학과
서울시 성북구 안암동 5가 1번지
Tel: 02-924-2054 Fax: 02-929-7914
Email: mjkim@nlp.korea.ac.kr
관심분야: 자연어처리, 기계번역

• 홍금원 (Hong, Gyunwon)

고려대학교 컴퓨터 · 전파통신공학과
서울시 성북구 안암동 5가 1번지
Tel: 02-924-2054 Fax: 02-929-7914
Email: gwhong@nlp.korea.ac.kr
관심분야: 자연어처리, 기계번역

• 박소영 (Park, So-Young)

상명대학교 디지털미디어학부
서울시 종로구 흥지동 7번지
Tel: 02-2287-5066 Fax: 02-2287-0072
Email: ssoya@smu.ac.kr
관심분야: 자연어처리

• 임해창 (Rim, Hae-Chang) 교신저자

고려대학교 컴퓨터 · 전파통신공학과
서울시 성북구 안암동 5가 1번지
Tel: 02-3290-3195 Fax: 02-929-7914
Email: rim@nlp.korea.ac.kr
관심분야: 자연어처리, 정보검색, 기계번역

부 록. 영어 음소를 대상으로 한 음자 표기 규칙

source	jamo	moc	moi	context	contjamo	end	target	koi	eni
a	mo	0	0	r(en,+1),\n(en,+2)	0	0	ㅏ(2,0)	0	+1
	mo	0	0	l(en,+1),l(en,+2)	0	0	ㅗ(2,0)	0	0
	mo	0	0	l(en,+1),t(en,+2)	0	0	ㅗ(2,0)	0	0
	mo	0	0	l(en,+1),k(en,+2)	0	0	ㅗ(2,0)	0	0
	mo	0	0	l(en,+1)	0	0	ㅏ(2,0)	0	0
	mo	0	0	u(en,+1)	0	0	ㅗ(2,0)	0	+1
	mo	0	0	w(en,+1)	0	0	ㅗ(2,0)	0	+1
	mo	0	0	y(en,+1)	0	0	ㅓ(2,0), ㅗ(1,+1), ㅏ(2,+2)	+2	+1
	mo	0	0	h(en,+1),\n(en,+2)	0	0	ㅏ(2,0)	0	+1
	mo	1	1	0	+ja	no	ㅐ(2,0)	0	0
	mo	1	1	0	0	0	ㅓ(2,0), ㅗ(1,+1), ㅏ(2,+2)	+2	0
	mo	2	0	0	+ja	0	ㅏ(2,0)	0	0
	mo	2	0	0	0	0	ㅓ(2,0), ㅗ(1,+1), ㅏ(2,+2)	+2	+1
e	mo	0	0	r(en,+1),\n(en,+2)	0	0	ㅓ(2,0)	0	+1
	mo	0	0	r(en,+1)	0	0	ㅓ(2,0)	0	0
	mo	2	2	0	+ja	0	ㅡ(2,0)	0	0
	mo	0	0	i(en,+1),g(en,g,+2),h(en,+3)	0	0	ㅓ(2,0), ㅗ(1,+1), ㅏ(2,+2)	+2	+3
	mo	0	0	i(en,+1)	0	0	ㅓ(2,0), ㅗ(1,+1), ㅏ(2,+2)	+2	+1
	mo	0	0	w(en,+1)	0	0	ㅠ(2,0)	0	+1
	mo	0	0	o(en,+1)	0	0	ㅓ(2,0)	0	+1
	mo	0	0	u(en,+1)	0	0	ㅠ(2,0)	0	+1
	mo	0	0	h(en,-1),t(en,-2)	0	0	ㅓ(2,0)	0	0
	mo	1	1	a(en,+1),r(en,+2)	0	0	ㅓ(2,0), ㅗ(1,+1), ㅓ(2,+2)	+2	+2
	mo	1	1	a(en,+1)	0	0	ㅓ(2,0)	0	+1
	mo	1	1	0	+ja	0	ㅓ(2,0)	0	0
	mo	1	1	0	0	0	ㅓ(2,0)	0	0
	mo	2	0	a(en,+1),r(en,+2)	0	0	ㅓ(2,0), ㅗ(1,+1), ㅓ(2,+2)	+2	+1
	mo	2	0	e(en,+1),r(en,+2)	0	0	ㅓ(2,0), ㅗ(1,+1), ㅓ(2,+2)	+2	+1
	mo	2	0	y(en,+1),r(en,+2)	0	0	ㅓ(2,0), ㅗ(1,+1), ㅓ(2,+2)	+2	+1
	mo	2	0	a(en,+1)	0	0	ㅓ(2,0)	0	+1
	mo	2	0	e(en,+1)	0	0	ㅓ(2,0)	0	+1
	mo	2	0	y(en,+1)	0	0	ㅓ(2,0)	0	+1
	mo	2	0	0	0	0	ㅓ(2,0)	0	0
i	mo	1	1	r(en,+1)	+mo	0	ㅓ(2,0), ㅗ(1,+1), ㅓ(2,+2)	+2	+1
	mo	2	0	r(en,+1)	+ja	0	ㅓ(2,0)	0	+1
	mo	0	0	n(en,+1),d(en,+2)	0	0	ㅓ(2,0), ㅗ(1,+1), ㅓ(2,+2)	+2	+1
	mo	0	0	l(en,+1),d(en,+2)	0	0	ㅓ(2,0), ㅗ(1,+1), ㅓ(2,+2)	+2	+1
	mo	0	0	e(en,+1)	0	0	ㅓ(2,0)	0	+1
	mo	0	0	i(en,+1)	0	0	ㅓ(2,0)	0	+1
	mo	0	0	a(en,+1),h(en,+2),\n(en,+3)	0	0	ㅓ(2,0), ㅗ(2,+1), ㅓ(2,+2)	+2	+2
	mo	0	0	a(en,+1)	0	0	ㅓ(2,0), ㅗ(2,+1), ㅓ(2,+2)	+2	+1
	mo	0	0	o(en,+1)	0	0	ㅓ(2,0), ㅗ(2,+1), ㅓ(2,+2)	+2	+1
	mo	0	0	u(en,+1)	0	0	ㅓ(2,0), ㅗ(2,+1), ㅓ(2,+2)	+2	+1
	mo	1	1	0	+ja	0	ㅓ(2,0)	0	0
	mo	1	1	0	0	0	ㅓ(2,0), ㅗ(1,+1), ㅓ(2,+2)	+2	0
	mo	2	0	0	0	0	ㅓ(2,0)	0	0
o	mo	0	0	o(en,+1)	0	0	ㅜ(2,0)	0	+1
	mo	0	0	u(en,+1)	0	0	ㅓ(2,0), ㅗ(1,+1), ㅜ(2,+2)	+2	+1
	mo	0	0	w(en,+1)	0	0	ㅓ(2,0), ㅗ(1,+1), ㅜ(2,+2)	+2	+1
	mo	0	0	y(en,+1)	0	0	ㅗ(2,0), ㅗ(1,+1), ㅓ(2,+2)	+2	+1
	mo	0	0	i(en,+1)	0	0	ㅗ(2,0), ㅗ(1,+1), ㅓ(2,+2)	+2	+1
	mo	0	0	a(en,+1)	0	0	ㅗ(2,0), ㅗ(1,+1), ㅓ(2,+2)	+2	+1
	mo	0	0	e(en,+1)	0	0	ㅗ(2,0), ㅗ(1,+1), ㅓ(2,+2)	+2	+1
	mo	0	0	0	0	0	ㅗ(2,0)	0	0
u	mo	0	0	r(en,+1)	+ja	0	ㅓ(2,0)	+1	+1

	mo	0	0	r(en,+1)	0	0	- (2,0)	0	+1
	mo	0	0	y(en,+1)	0	0	↑ (2,0), o (1,+1), l (2,+2)	+2	+1
	mo	1	1	0	+ja	0	↑ (2,0)	0	0
	mo	1	1	0	0	0	π (2,0)	0	0
	mo	2	0	0	0	0	π (2,0)	0	0
	mo	0	0	0	0	0	π (2,0)	0	0
w	mo	0	0	a(en,+1),r(en,+2)	0	0	¬ (2,0)	0	+2
	mo	0	0	a(en,+1)	0	0	↑ (2,0)	0	+1
	mo	0	0	e(en,+1),a(en,+2)	0	0	¬ (2,0)	0	+2
	mo	0	0	e(en,+1)	0	0	¬ (2,0)	0	+1
	mo	0	0	i(en,+1)	0	0	¬ (2,0)	0	+1
	mo	0	0	o(en,+1)	0	0	¬ (2,0)	0	+1
	mo	0	0	u(en,+1)	0	0	π (2,0)	0	+1
	mo	0	0	h(en,+1),e(en,+2)	0	0	¬ (2,0)	0	+2
	mo	0	0	h(en,+1),a(en,+2)	0	0	¬ (2,0)	0	+2
	mo	0	0	h(en,+1),i(en,+2)	0	0	¬ (2,0)	0	+2
	mo	0	0	h(en,+1),o(en,+2)	0	0	¬ (1,-1), π (2,0)	+1	+2
	mo	0	0	h(en,+1),y(en,+2)	0	0	¬ (2,0), o (1,+1), l (2,+2)	+2	+2
	mo	0	0	0	0	0	π (2,0)	0	0
y	mo	0	0	a(en,+1),e(en,+2)	0	0	¬ (2,0)	0	+2
	mo	0	0	a(en,+1)	0	0	¬ (2,0)	0	+1
	mo	0	0	e(en,+1),o(en,+2)	0	0	¬ (2,0)	0	+1
	mo	0	0	e(en,+1)	0	0	¬ (2,0)	0	+1
	mo	0	0	i(en,+1)	0	0	l (2,0)	0	+1
	mo	0	0	o(en,+1),u(en,+2),n(en,+3),g(en,+4)	0	0	¬ (2,0)	0	+2
	mo	0	0	o(en,+1),u(en,+2)	0	0	π (2,0)	0	+2
	mo	0	0	o(en,+1)	0	0	¬ (2,0)	0	+1
	mo	0	0	u(en,+1)	0	0	π (2,0)	0	+1
	mo	2	2	0	0	0	l (2,0)	0	0
	mo	1	1	0	0	0	¬ (2,0), o (1,+1), l (2,+2)	+2	0
	mo	0	0	0	0	0	l (2,0)	0	0
b	ja	0	0	0	-ja	yes	¬ (1,0), -(2,+1)	+1	0
	ja	0	0	0	-ja,+ja	0	¬ (1,0), -(2,+1)	+1	0
	ja	0	0	0	-ja,+mo	0	¬ (1,0)	0	0
	ja	0	0	0	+mo	0	¬ (1,0)	0	0
	ja	0	0	0	-mo	0	¬ (3,0)	0	0
	ja	0	0	0	0	0	¬ (1,0), -(2,+1)	+1	0
c	ja	0	0	e(en,+1)	0	0	¬ (1,0)	0	0
	ja	0	0	i(en,+1)	0	0	¬ (1,0)	0	0
	ja	0	0	y(en,+1)	0	0	¬ (1,0)	0	0
	ja	0	0	a(en,+1)	0	0	¬ (1,0)	0	0
	ja	0	0	o(en,+1)	0	0	¬ (1,0)	0	0
	ja	0	0	u(en,+1)	0	0	¬ (1,0)	0	0
	ja	0	0	h(en,+1),e(en,+2)	0	0	¬ (1,0)	0	+1
	ja	0	0	h(en,+1),i(en,+2)	0	0	¬ (1,0)	0	+1
	ja	0	0	h(en,+1),y(en,+2)	0	0	¬ (1,0)	0	+1
	ja	0	0	h(en,+1),o(en,+2)	0	0	¬ (1,0)	0	+1
	ja	0	0	h(en,+1),n(en,+2)	0	0	¬ (1,0), -(2,+1)	+1	+1
	ja	0	0	h(en,+1),a(en,+2)	0	0	¬ (1,0)	0	+1
	ja	0	0	h(en,+1),w(en,+2),s(en,-1)	0	0	o (1,0)	0	+1
	ja	0	0	h(en,+1)	+ja	0	¬ (1,0), -(2,+1)	+1	+1
	ja	0	0	k(en,+1),n(en,+2)	-mo	0	¬ (3,0)	0	+1
	ja	0	0	k(en,+1),n(en,+2)	0	0	¬ (1,0)	0	+1
	ja	0	0	k(en,+1)	+ja,-mo	0	¬ (3,0)	0	+1
	ja	0	0	k(en,+1)	+ja	0	¬ (1,0)	0	+1
	ja	0	0	k(en,+1)	0	0	¬ (1,0)	0	+1

	ja	0	0	c(en,+1)	+ja	0	$\neg(1,0), -(2,+1)$	+1	+1
	ja	0	0	0	-mo	yes	$\neg(3,0)$	0	0
	ja	0	0	0	0	yes	$\neg(1,0), -(2,+1)$	+1	0
d	ja	0	0	g(en,+1),e(en,+2)	0	0	$\neg(1,0), \neg(2,+1)$	+1	+2
	ja	0	0	0	0	yes	$\neg(1,0), -(2,+1)$	+1	0
	ja	0	0	0	+ja	0	$\neg(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	0	$\neg(1,0)$	0	0
f	ja	0	0	f(en,+1)	+ja	0	$\neg(1,0), -(2,+1)$	+1	+1
	ja	0	0	f(en,+1)	0	yes	$\neg(1,0), -(2,+1)$	+1	+1
	ja	0	0	0	+ja	0	$\neg(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	yes	$\neg(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	0	$\neg(1,0)$	0	0
g	ja	0	0	h(en,+1)	+ja	0	$\neg(1,0), -(2,+1)$	+1	+1
	ja	0	0	h(en,+1)	0	yes	$\neg(1,0), -(2,+1)$	+1	+1
	ja	0	0	h(en,+1)	0	0	$\neg(1,0)$	0	+1
	ja	0	0	n(en,+1)	0	yes	$\neg(3,0)$	0	+1
	ja	0	0	n(en,+1)	0	0	$\neg(1,0)$	0	+1
	ja	0	0	e(en,+1)	0	0	$\neg(1,0)$	0	0
	ja	0	0	i(en,+1)	0	0	$\neg(1,0)$	0	0
	ja	0	0	y(en,+1)	0	0	$\neg(1,0)$	0	0
	ja	0	0	a(en,+1)	0	0	$\neg(1,0)$	0	0
	ja	0	0	o(en,+1)	0	0	$\neg(1,0)$	0	0
	ja	0	0	u(en,+1)	0	0	$\neg(1,0)$	0	0
	ja	0	0	w(en,+1)	0	0	$\neg(1,0)$	0	0
	ja	0	0	r(en,+1)	+ja	0	$\neg(1,0), -(2,+1)$	+1	+1
	ja	0	0	r(en,+1)	0	yes	$\neg(1,0), -(2,+1)$	+1	+1
h	ja	0	0	n(en,+1)	0	0	$\neg(1,0), \neg(2,+1), \neg(3,+2)$	+2	+1
	ja	0	0	l(en,+1)	-mo	0	$\neg(3,0)$	0	+1
	ja	0	0	0	+ja	0	$\neg(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	yes	$\neg(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	0	$\neg(1,0)$	0	0
j	ja	0	0	0	+ja	0	$\neg(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	yes	$\neg(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	0	$\neg(1,0)$	0	0
k	ja	0	0	n(en,+1)	0	0	$\neg(1,0)$	0	+1
	ja	0	0	0	+ja	0	$\neg(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	yes	$\neg(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	0	$\neg(1,0)$	0	0
l	ja	0	0	l(en,+1)	-mo,+mo	0	$\neg(3,0) \neg(1,+1)$	+1	+1
	ja	0	0	m(en,+1)	-mo,+mo	0	$\neg(3,0) \neg(1,+1)$	+1	0
	ja	0	0	0	-mo,+mo	0	$\neg(3,0) \neg(1,+1)$	+1	0
	ja	0	0	0	+ja	0	$\neg(1,0), -(2,+1)$	0	0
	ja	0	0	0	0	yes	$\neg(1,0), -(2,+1)$	0	0
	ja	0	0	0	0	0	$\neg(1,0)$	0	0
m	ja	0	0	m(en,+1)	-mo,+mo	no	$\neg(3,0), \neg(1,+1)$	+1	+1
	ja	0	0	m(en,+1)	-mo	0	$\neg(3,0)$	0	+1
	ja	0	0	m(en,+1)	0	0	$\neg(1,0)$	0	+1
	ja	0	0	0	+ja	0	$\neg(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	yes	$\neg(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	0	$\neg(1,0)$	0	0
n	ja	0	0	g(en,+1),e(en,+2)	0	0	$\circ(3,0), \neg(1,+1)$	+1	+1
	ja	0	0	g(en,+1),i(en,+2)	0	0	$\circ(3,0), \neg(1,+1)$	+1	+1
	ja	0	0	g(en,+1),y(en,+2)	0	0	$\circ(3,0), \neg(1,+1)$	+1	+1
	ja	0	0	g(en,+1),a(en,+2)	0	0	$\circ(3,0), \neg(1,+1)$	+1	+1
	ja	0	0	g(en,+1),o(en,+2)	0	0	$\circ(3,0), \neg(1,+1)$	+1	+1
	ja	0	0	g(en,+1),w(en,+2)	0	0	$\circ(3,0), \neg(1,+1)$	+1	+1
	ja	0	0	g(en,+1),u(en,+2)	0	0	$\circ(3,0), \neg(1,+1)$	+1	+1

	ja	0	0	$g(en,+1)$	0	0	$\circ(3,0)$	0	+1
	ja	0	0	$k(en,+1)$	+ja	0	$\circ(3,0), \neg(1,+1), -(2,+2)$	+2	+1
p	ja	0	0	0	+mo	0	$\boxdot(1,0)$	0	0
	ja	0	0	$h(en,+1)$	+ja	0	$\boxminus(1,0), -(2,+1)$	+1	+1
	ja	0	0	$h(en,+1)$	0	yes	$\boxminus(1,0), -(2,+1)$	+1	+1
	ja	0	0	$h(en,+1)$	0	0	$\boxminus(1,0)$	0	+1
	ja	0	0	0	+ja	0	$\boxminus(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	yes	$\boxminus(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	0	$\boxminus(1,0)$	0	0
	q	ja	0	0	0	0	$\neg(1,0)$	0	0
r	ja	0	0	0	0	yes	$\circ(1,0), \dashv(2,+1)$	+1	0
	ja	0	0	0	+mo	0	$\boxdot(1,0)$	0	0
	ja	0	0	$m(en,+1)$	-mo, +mo	0	$\square(1,0)$	0	+1
	ja	0	0	$m(en,+1)$	-mo	0	$\square(3,0)$	0	+1
	ja	0	0	0	0	yes	$\boxminus(1,0), -(2,+1)$	+1	+1
	ja	0	0	$h(en,+1)$	+mo	0	$\boxminus(1,0)$	0	+1
	ja	0	0	$h(en,+1)$	0	yes	$\boxminus(1,0), -(2,+1)$	+1	+1
	ja	0	0	$r(en,+1)$	-mo	yes	0	0	+1
	ja	0	0	$r(en,+1)$	-ja	yes	$\boxminus(1,0), -(2,+1)$	+1	+1
	ja	0	0	$r(en,+1)$	+mo	0	$\boxminus(1,0)$	0	+1
	ja	0	0	$r(en,+1), h(en,+2)$	+mo	0	$\boxminus(1,0)$	0	+2
	ja	0	0	$r(en,+1), h(en,+2)$	0	0	$\boxminus(1,0), -(2,+1)$	+1	+2
	ja	0	0	$r(en,+1)$	+ja	0	$\boxminus(1,0), -(2,+1)$	+1	+1
	ja	0	0	0	-mo, +ja	0	0	0	+1
	ja	0	0	0	-ja, +ja	0	$\circ(1,0), \dashv(2,+1)$	+1	+1
s	ja	0	0	$s(en,+1), i(en,+2), o(en,+3), n(en,+4)$	0	0	$\wedge(1,0), \exists(2,+1), \sqcup(3,+2)$	+2	+4
	ja	0	0	$i(en,+1), o(en,+2), n(en,+3)$	0	0	$\wedge(1,0), \exists(2,+1), \sqcup(3,+2)$	+2	+3
	ja	0	0	$s(en,+1)$	-mo, +mo	0	$\boxtimes(1,0)$	0	+1
	ja	0	0	$c(en,+1), h(en,+2)$	0	0	$\wedge(1,0), -(2,+1)$	+1	+2
	ja	0	0	$s(en,+1), m(en,+2)$	0	0	$-(2,0), \square(3,+1)$	0	+2
	ja	0	0	$c(en,+1)$	+ja	0	$\wedge(1,0), -(2,+1)$	+1	+1
	ja	0	0	$c(en,+1)$	0	yes	$\wedge(1,0), -(2,+1)$	+1	+1
	ja	0	0	$c(en,+1)$	0	0	$\wedge(1,0)$	0	+1
	ja	0	0	$h(en,+1)$	+ja	0	$\wedge(1,0), -(2,+1)$	+1	+1
	ja	0	0	$h(en,+1)$	0	yes	$\wedge(1,0), -(2,+1)$	+1	+1
	ja	0	0	$h(en,+1)$	0	0	$\wedge(1,0)$	0	+1
	ja	0	0	$s(en,+1)$	+ja	0	$\wedge(1,0), -(2,+1)$	+1	+1
	ja	0	0	$s(en,+1)$	0	yes	$\wedge(1,0), -(2,+1)$	+1	+1
	ja	0	0	0	0	0	$\wedge(1,0), -(2,+1)$	+1	0
t	ja	0	0	$i(en,+1), o(en,+2), n(en,+3)$	0	0	$\wedge(1,0), \exists(2,+1), \sqcup(3,+2)$	+2	+3
	ja	0	0	0	+mo	0	$\boxminus(1,0)$	0	0
	ja	0	0	$h(en,+1)$	+ja	0	$\sqsubseteq(1,0), -(2,+1)$	+1	+1
	ja	0	0	$h(en,+1)$	0	yes	$\sqsubseteq(1,0), -(2,+1)$	+1	+1
	ja	0	0	$h(en,+1)$	0	0	$\sqsubseteq(1,0)$	0	+1
	ja	0	0	0	+ja	0	$\boxminus(1,0), -(2,+1)$	+1	0
v	ja	0	0	0	+ja	0	$\boxdot(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	yes	$\boxdot(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	0	$\boxdot(1,0)$	0	0
x	ja	0	0	0	-mo	0	$\neg(3,0)$	0	0
	ja	0	0	0	+ja	0	$\wedge(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	yes	$\wedge(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	0	$\wedge(1,0)$	0	0
z	ja	0	0	0	+ja	0	$\boxtimes(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	yes	$\boxtimes(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	0	$\boxtimes(1,0)$	0	0