



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2011-0091294
(43) 공개일자 2011년08월11일

(51) Int. Cl.

G06F 17/28 (2006.01) G06F 17/27 (2006.01)

(21) 출원번호 10-2010-0011061

(22) 출원일자 2010년02월05일

심사청구일자 2010년02월05일

(71) 출원인

고려대학교 산학협력단

서울 성북구 안암동5가 1

(72) 발명자

임해창

서울 서초구 서초1동 1642-34 래미안아파트 서초 7차 102-1002

김민정

서울시 용산구 이촌1동 장미맨션 6층 D호

(74) 대리인

특허법인남춘

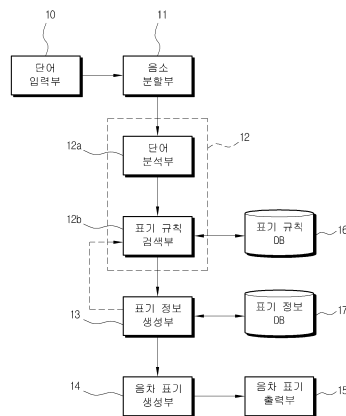
전체 청구항 수 : 총 14 항

(54) 외래어의 자동 음차 표기 방법 및 장치

(57) 요약

본 발명은 외래어의 자동 음차 표기 방법 및 장치에 관한 것이다. 본 발명에 따른 자동 음차 표기 방법은 (a) 외래어를 구성하는 각 음소에 대한 타겟 음소 정보와, 상기 각 음소의 주변 음소에 대한 주변 음소 정보에 기초하여, 각 음소별로 복수의 음소 단위 음차 표기 규칙을 생성하는 단계와; (b) 변환 대상 단어를 음소 단위로 분할하는 단계와; (c) 상기 복수의 음소 단위 음차 표기 규칙으로부터 상기 분할된 각 음소에 해당하는 음소 단위 음차 표기 규칙을 추출하는 단계와; (d) 상기 추출된 음소 단위 음차 표기 규칙에 따라 상기 분할된 각 음소에 대한 음차 표기 정보를 생성하는 단계와; (e) 상기 생성된 음차 표기 정보에 기초하여 상기 변환 대상 단어에 대한 음차 표기를 생성하는 단계를 포함하는 것을 특징으로 한다. 이에 따라, 음철법(Phonics)의 개념을 이용하여 음차 표기 규칙을 생성하여 이를 자동 음차 표기에 적용함으로써, 보다 정확하고 용례를 많이 수집하기 어려운 경우나 신조어 등의 처리에도 강인하게 음차 표기가 가능하게 된다.

대표도 - 도1



특허청구의 범위

청구항 1

외래어의 자동 음차 표기 방법에 있어서,

- (a) 외래어를 구성하는 각 음소에 대한 타겟 음소 정보와, 상기 각 음소의 주변 음소에 대한 주변 음소 정보에 기초하여, 각 음소별로 복수의 음소 단위 음차 표기 규칙을 생성하는 단계와;
- (b) 변환 대상 단어를 음소 단위로 분할하는 단계와;
- (c) 상기 복수의 음소 단위 음차 표기 규칙으로부터 상기 분할된 각 음소에 해당하는 음소 단위 음차 표기 규칙을 추출하는 단계와;
- (d) 상기 추출된 음소 단위 음차 표기 규칙에 따라 상기 분할된 각 음소에 대한 음차 표기 정보를 생성하는 단계와;
- (e) 상기 생성된 음차 표기 정보에 기초하여 상기 변환 대상 단어에 대한 음차 표기를 생성하는 단계를 포함하는 것을 특징으로 하는 외래어의 자동 음차 표기 방법.

청구항 2

제1항에 있어서,

상기 타겟 음소 정보는,

상기 타겟 음소 정보의 대상이 되는 타겟 음소와,

상기 타겟 음소가 자음인지 모음인지에 대한 타겟 자모 정보와,

단어 내의 모음의 개수에 대한 모음 개수 정보와,

상기 타겟 음소가 단어 내에서 몇 번째 모음인지에 대한 모음 위치 정보를 포함하며;

상기 주변 음소 정보는,

상기 주변 음소 정보에 적용될 대상이 되는 타겟 주변 음소와, 상기 타겟 음소와 상기 타겟 주변 음소 간의 거리에 대한 주변 음소 거리 정보를 갖는 콘텍스트 정보와,

상기 타겟 주변 음소가 자음인지 모음인지에 대한 주변 자모 정보와,

상기 타겟 음소가 단어의 마지막 음소인지 여부에 대한 단어 말단 정보를 포함하는 것을 특징으로 하는 외래어의 자동 음차 표기 방법.

청구항 3

제2항에 있어서,

음소별로 생성되는 상기 복수의 음소 단위 음차 표기 규칙은 음소별로 기 설정된 우선 순위를 가지며;

상기 (c) 단계는,

- (c1) 상기 추출된 음소가 상기 복수의 음소 단위 음차 표기 규칙 중 어느 하나에 부합하는지 여부를 상기 우선 순위에 따라 순차적으로 체크하는 단계와,
- (c2) 상기 (c1) 단계의 수행 과정 중 상기 추출된 음소가 부합되는 최우선 순위의 어느 하나를 상기 분할된 음소에 해당하는 상기 음소 단위 음차 표기 규칙으로 추출하는 단계를 포함하는 것을 특징으로 하는 외래어의 자동 음차 표기 방법.

청구항 4

제3항에 있어서,

상기 타겟 주변 음소는 외래어 표기시 상기 타겟 음소의 주변에 표기되어 상기 타겟 음소의 발음을 가변시키는 음소로 설정되는 것을 특징으로 하는 외래어의 자동 음차 표기 방법.

청구항 5

제3항에 있어서,

상기 각 음소 단위 음차 표기 규칙은 고려 대상 정보와 비고려 대상 정보에 기초하여 생성되며;

상기 모음 개수 정보, 상기 모음 위치 정보, 상기 콘텍스트 정보, 상기 주변 자모 정보 및 상기 단어 말단 정보는 음소별 및 음소 단위 음차 표기 규칙별로 상기 고려 대상 정보 또는 상기 비고려 대상 정보에 선택적으로 적용되는 것을 특징으로 하는 외래어의 자동 음차 표기 방법.

청구항 6

제5항에 있어서,

상기 모음 개수 정보는 단어 내에서 모음의 개수가 하나 또는 둘 이상인지 여부에 따라 이진 분류되는 것을 특징으로 하는 외래어의 자동 음차 표기 방법.

청구항 7

제3항에 있어서,

상기 (d) 단계에서 생성되는 상기 음차 표기 정보는,

상기 분할된 음소에 대해 표기될 음차 표기에 대한 적어도 하나의 단위 표기 정보와;

상기 각 단위 표기 정보가 초성, 중성 및 종성 중 어디에 속하는지 여부에 대한 삼성 정보와;

현재 생성되는 음차 표기 정보의 대상이 되는 음소 이후의 음소 중 상기 (c) 단계의 수행을 스킵(skip)할 적어도 하나의 음소에 대한 스킵 음소 정보를 포함하는 것을 특징으로 하는 외래어의 자동 음차 표기 방법.

청구항 8

외래어의 자동 음차 표기 장치에 있어서,

외래어를 구성하는 각 음소에 대한 타겟 음소 정보와, 상기 각 음소의 주변 음소에 대한 주변 음소 정보에 기초하여 생성되며 각 음소별로 복수개로 마련되는 음소 단위 음차 표기 규칙이 저장된 표기 규칙 데이터베이스와;

상기 각 음소 단위 음차 표기 규칙에 대응하도록 마련된 음차 표기 정보가 저장된 표기 정보 데이터베이스와;

변환 대상 단어를 음소 단위로 분할하는 음소 분할부와;

상기 표기 규칙 데이터베이스로부터 상기 음소 분할부에 의해 분할된 각 음소에 대응하는 음소 단위 음차 표기 규칙을 추출하는 표기 규칙 추출부와;

상기 표기 규칙 추출부에 의해 추출된 상기 음소 단위 음차 표기 규칙에 대응하는 음차 표기 정보를 상기 표기 정보 데이터베이스로부터 추출하여 상기 각 음소에 대한 음차 표기 정보를 생성하는 표기 정보 생성부와;

상기 표기 정보 생성부에 의해 생성된 상기 음차 표기 정보에 기초하여 상기 변환 대상 단어에 대한 음차 표기를 생성하는 음차 표기 생성부를 포함하는 것을 특징으로 하는 외래어의 자동 음차 표기 장치.

청구항 9

제8항에 있어서,

상기 타겟 음소 정보는,

상기 타겟 음소 정보의 대상이 되는 타겟 음소와,

상기 타겟 음소가 자음인지 모음인지에 대한 타겟 자모 정보와,

단어 내의 모음의 개수에 대한 모음 개수 정보와,

상기 타겟 음소가 단어 내에서 몇 번째 모음인지에 대한 모음 위치 정보를 포함하며;

상기 주변 음소 정보는,

상기 주변 음소 정보에 적용될 대상이 되는 타겟 주변 음소와, 상기 타겟 음소와 상기 타겟 주변 음소 간의 거리에 대한 주변 음소 거리 정보를 갖는 콘텍스트 정보와,

상기 타겟 주변 음소가 자음인지 모음인지에 대한 주변 자모 정보와,

상기 타겟 음소가 단어의 마지막 음소인지 여부에 대한 단어 말단 정보를 포함하는 것을 특징으로 하는 외래어의 자동 음차 표기 장치.

청구항 10

제9항에 있어서,

상기 표기 규칙 추출부는,

상기 추출된 음소가 상기 복수의 음소 단위 음차 표기 규칙 중 어느 하나에 부합되는지 여부를 각 음소별로 상기 복수의 음소 단위 음차 표기 규칙에 대해 기 설정된 우선 순위에 따라 순차적으로 체크하고;

상기 복수의 음소 단위 음차 표기 규칙 중 상기 추출된 음소가 부합되는 최우선 순위의 어느 하나를 상기 분할된 음소에 해당하는 상기 음소 단위 음차 표기 규칙으로 추출하는 것을 특징으로 하는 외래어의 자동 음차 표기 장치.

청구항 11

제10항에 있어서,

상기 타겟 주변 음소는 외래어 표기시 상기 타겟 음소의 주변에 표기되어 상기 타겟 음소의 발음을 가변시키는 음소로 설정되는 것을 특징으로 하는 외래어의 자동 음차 표기 장치.

청구항 12

제10항에 있어서,

상기 각 음소 단위 음차 표기 규칙은 고려 대상 정보와 비고려 대상 정보에 기초하여 생성되며;

상기 모음 개수 정보, 상기 모음 위치 정보, 상기 콘텍스트 정보, 상기 주변 자모 정보 및 상기 단어 말단 정보는 음소별 및 음소 단위 음차 표기 규칙별로 상기 고려 대상 정보 또는 상기 비고려 대상 정보에 선택적으로 적용되는 것을 특징으로 하는 외래어의 자동 음차 표기 장치.

청구항 13

제12항에 있어서,

상기 모음 개수 정보는 단어 내에서 모음의 개수가 하나 및 둘 이상인지 여부에 따라 이진 분류되는 것을 특징으로 하는 외래어의 자동 음차 표기 장치.

청구항 14

제10항에 있어서,

상기 음차 표기 정보는,

상기 분할된 음소의 음차 표기 방식에 대한 적어도 하나의 단위 표기 정보와;

상기 각 단위 표기 정보가 초성, 중성 및 종성 중 어디에 속하는지 여부에 대한 삼성 정보와;

현재 생성되는 음차 표기 정보의 대상이 되는 음소 이후의 음소 중 상기 표기 규칙 추출부가 상기 음소 단위 음차 표기 규칙의 추출을 스킵(skip)할 적어도 하나의 음소에 대한 스킵 음소 정보를 포함하는 것을 특징으로 하는 외래어의 자동 음차 표기 장치.

명세서

기술분야

[0001] 본 발명은 외래어의 자동 음차 표기 방법 및 장치에 관한 것으로서, 보다 상세하게는 음철법(Phonics)의 개념을 이용하여 보다 정확한 음차 표기가 가능한 외래어의 자동 음차 표기 방법 및 장치에 관한 것이다.

배경기술

[0002] 외국과의 교류가 빈번해지면서 신문 기사를 비롯한 한국어 문서에서 많은 외국어 단어(이하, '외래어'라 함)들을 발음에 따라 한글로 표기하고 있다. 이와 같이, 외래어의 소리에 따라 문자나 말을 바꿔 쓰는 것을 음차(Transliteration)라고 한다.

[0003] 오래전부터 한국어는 한자어들을 음차 표기하여 사용하여 왔으며, 최근에는 중국어, 일본어, 한국어에서 영어, 프랑스어, 독일어 등의 단어들을 음차 표기하고 있다. 일 예로, 텔레비전, 디지털 카메라와 같은 일반 명사들과 밴쿠버, 샌프란시스코, 헬렌 켈러와 같은 인명과 지명을 포함하는 고유명사들도 음차 표기하여 쓴다.

[0004] 이와 같은, 음차 표기의 결과물은 사회 전반에 걸쳐 다양한 용도, 예컨대 자연어 처리 분야에서 널리 사용되고 있다. 또한, 음차 표기의 다른 이용 분야의 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 구축을 위한 분야에도 적용되고 있으며, 이외에도 기계 번역 등의 분야에서 사용이 가능하다.

[0005] 이와 같은 음차 표기와 관련하여, 특정 단어가 입력되면 이를 자동으로 음차 표기로 변환하는 기술이 제안되고 있으며, 이러한 자동 음차 표기 기술에 접근하는 방법으로 사전 기반 음차 표기 방법, 통계 기반 음차 표기 방법, 규칙 기반 음차 표기 방법이 제안되고 있다.

[0006] 사전 기반 음차 표기 방법은 원어와 음차 표기한 단어의 쌍이 있는 사전에서 가장 비슷한 원어 단어를 찾고 그에 대응되는 음차 표기한 단어를 이용하는 방법이다. 이와 같은 사전 기반 음차 표기 방법에 따라 음차 표기를 수행할 수 있었을 때, 즉 해당 단어와 음차 표기가 사전에 등록되어 있을 때의 결과는 매우 정확한 편이어서 자연스러운 음차 표기가 가능하나, 실제로 실생활에서 사용하는 모든 단어를 사전으로 구축하는 것은 불가능하여 미등록 단어에 대해서는 정확한 음차 표기가 어려운 단점이 있다.

[0007] 통계 기반 음차 표기 방법은 외래어와 음차 표기한 한국어 쌍으로 구성된 말뭉치를 이용하는 방법으로, 최대 엔트로피 모델을 이용한 방법, 결정 트리를 이용한 방법, 메모리 기반 학습과 결정 트리를 사용한 방법, 음운 패턴을 이용한 방법, 음소 단위 n-gram 음차 표기 확률을 이용한 방법 등 다양한 형태의 음차 표기 방법이 제안되고 있다.

[0008] 통계 기반 음차 표기 방법은 사전 기반 음차 표기 방법에 비해 자료 부족 문제에 매우 강건한 장점을 가지고 있다. 그러나, 단순히 통계 자료만으로 음차 표기를 하다보면 학습에 사용한 데이터와 실제 음차 표기를 수행해야 하는 단어의 성격이 다를 경우 전혀 엉뚱하게 발음되는 음차 표기가 발생할 경우가 많다. 예를 들어, 인명, 지명 표기 용례로 학습한 통계 기반 음차 표기 방법으로 일반 단어의 음차 표기를 수행했을 때 'Sunday'를 '순데이'로, 'Sale'을 '살레'로 음차 표기하는 오류를 발생시킬 수 있게 된다. 자연어 처리 분야에서 음차 표기 결과물을 사용할 때에는 어느 정도의 오류가 있어도 사용이 가능하나, 외국어 교육이나 문화, 관광과 같은 분야에 음차 표기를 사용하는 경우에는 어느 정도의 비용을 감수하고 사람이 개입하여 음차 표기 성능을 개선하여야 한다.

[0009] 따라서, 통계 기반 음차 표기 방법에서 오류를 최소화하기 위해서는 음차 표기의 예를 포함하는 좋은 품질의 말뭉치를 다량으로 수집 혹은 구축하는 것이 필요하며, 이는 많은 시간과 비용을 요구하게 되는 단점이 있다. 또한, 통계 기반 음차 표기 방법은 상대적으로 용례 수집이 어려운 비영어권 단어나 신조어의 처리에 취약한 단점이 있다.

[0010] 규칙 기반 음차 표기 방법은 사람이 수동으로 구축한 규칙이나 말뭉치에서 자동으로 추출한 규칙을 이용하여 음차 표기를 수행하는 방법이다. 규칙 기반 음차 표기 방법은 규칙을 구축하는데 비용과 시간이 다소 소요되는 단점이 있으나, 용례를 많이 수집하기 어려운 경우나 신조어 등의 처리에 있어 사전 기반 음차 표기 방법이나 통계 기반 음차 표기 방법에 비해 강인할 수 있다.

[0011] 외래어를 한국어로 음차 표기하는데 사용되는 기존의 규칙 기반 음차 표기 방법은 외래어 표기법을 사용하고 있

다. 여기서, 외래어 표기법을 규칙 기반 음차 표기 방법에 적용하기 위해서는 해당 외래어를 발음으로 전환하고, 이를 다시 한국어로 음차 표기하는 두 번의 과정이 필요하게 된다.

[0012] 그런데, 외래어 표기법을 사용함에 따른 두 번의 변환 과정에서, 첫 번째 단계, 즉 외래어를 발음으로 전환하는 단계에서 오류가 발생하는 경우 이러한 오류가 음차 표기로까지 전파되어 전체적인 음차 표기 성능을 하락시키는 문제점이 있다.

발명의 내용

해결하려는 과제

[0013] 본 발명은 상기와 같은 기존의 음차 표기 방법, 즉 상술한 사전 기반 음차 표기 방법, 통계 기반 음차 표기 방법, 그리고 규칙 기반 음차 표기 방법이 갖는 문제점을 해결하기 위해 안출된 것으로서, 음철법(Phonics)의 개념을 이용하여 음차 표기 규칙을 생성하여 이를 자동 음차 표기에 적용함으로써, 보다 정확하고 용례를 많이 수집하기 어려운 경우나 신조어 등의 처리에도 강인한 외래어의 자동 음차 표기 방법 및 장치를 제공하는데 그 목적이 있다.

과제의 해결 수단

[0014] 상기 목적은 본 발명에 따라, 외래어의 자동 음차 표기 방법에 있어서, (a) 외래어를 구성하는 각 음소에 대한 타겟 음소 정보와, 상기 각 음소의 주변 음소에 대한 주변 음소 정보에 기초하여, 각 음소별로 복수의 음소 단위 음차 표기 규칙을 생성하는 단계와; (b) 변환 대상 단어를 음소 단위로 분할하는 단계와; (c) 상기 복수의 음소 단위 음차 표기 규칙으로부터 상기 분할된 각 음소에 해당하는 음소 단위 음차 표기 규칙을 추출하는 단계와; (d) 상기 추출된 음소 단위 음차 표기 규칙에 따라 상기 분할된 각 음소에 대한 음차 표기 정보를 생성하는 단계와; (e) 상기 생성된 음차 표기 정보에 기초하여 상기 변환 대상 단어에 대한 음차 표기를 생성하는 단계를 포함하는 것을 특징으로 하는 외래어의 자동 음차 표기 방법에 의해서 달성될 수 있다.

[0015] 또한, 상기 목적은 본 발명의 다른 실시 형태에 따라, 외래어의 자동 음차 표기 장치에 있어서, 외래어를 구성하는 각 음소에 대한 타겟 음소 정보와, 상기 각 음소의 주변 음소에 대한 주변 음소 정보에 기초하여 생성되되 각 음소별로 복수개로 마련되는 음소 단위 음차 표기 규칙이 저장된 표기 규칙 데이터베이스와; 상기 각 음소 단위 음차 표기 규칙에 대응하도록 마련된 음차 표기 정보가 저장된 표기 정보 데이터베이스와; 변환 대상 단어를 음소 단위로 분할하는 음소 분할부와; 상기 표기 규칙 데이터베이스로부터 상기 음소 분할부에 의해 분할된 각 음소에 대응하는 음소 단위 음차 표기 규칙을 추출하는 표기 규칙 추출부와; 상기 표기 규칙 추출부에 의해 추출된 상기 음소 단위 음차 표기 규칙에 대응하는 음차 표기 정보를 상기 표기 정보 데이터베이스로부터 추출하여 상기 각 음소에 대한 음차 표기 정보를 생성하는 표기 정보 생성부와; 상기 표기 정보 생성부에 의해 생성된 상기 음차 표기 정보에 기초하여 상기 변환 대상 단어에 대한 음차 표기를 생성하는 음차 표기 생성부를 포함하는 것을 특징으로 하는 외래어의 자동 음차 표기 장치에 의해서도 달성될 수 있다.

[0016] 여기서, 상기 타겟 음소 정보는 상기 타겟 음소 정보의 대상이 되는 타겟 음소와, 상기 타겟 음소가 자음인지 모음인지에 대한 타겟 자모 정보와, 단어 내의 모음의 개수에 대한 모음 개수 정보와, 상기 타겟 음소가 단어 내에서 몇 번째 모음인지에 대한 모음 위치 정보를 포함하며; 상기 주변 음소 정보는 상기 주변 음소 정보에 적용될 대상이 되는 타겟 주변 음소와, 상기 타겟 음소와 상기 타겟 주변 음소 간의 거리에 대한 주변 음소 거리 정보를 갖는 콘텍스트 정보와, 상기 타겟 주변 음소가 자음인지 모음인지에 대한 주변 자모 정보와, 상기 타겟 음소가 단어의 마지막 음소인지 여부에 대한 단어 말단 정보를 포함할 수 있다.

[0017] 그리고, 음소별로 생성되는 상기 복수의 음소 단위 음차 표기 규칙은 음소별로 기 설정된 우선 순위를 가지며; 상기 (c) 단계는 (c1) 상기 추출된 음소가 상기 복수의 음소 단위 음차 표기 규칙 중 어느 하나에 부합하는지 여부를 상기 우선 순위에 따라 순차적으로 체크하는 단계와, (c2) 상기 (c1) 단계의 수행 과정 중 상기 추출된 음소가 부합되는 최우선 순위의 어느 하나를 상기 분할된 음소에 해당하는 상기 음소 단위 음차 표기 규칙으로 추출하는 단계를 포함할 수 있다.

[0018] 그리고, 상기 타겟 주변 음소는 외래어 표기시 상기 타겟 음소의 주변에 표기되어 상기 타겟 음소의 발음을 가변시키는 음소로 설정될 수 있다.

[0019] 또한, 상기 각 음소 단위 음차 표기 규칙은 고려 대상 정보와 비고려 대상 정보에 기초하여 생성되며; 상기 모

음 개수 정보, 상기 모음 위치 정보, 상기 콘텍스트 정보, 상기 주변 자모 정보 및 상기 단어 말단 정보는 음소 별 및 음소 단위 음차 표기 규칙별로 상기 고려 대상 정보 또는 상기 비고려 대상 정보에 선택적으로 적용될 수 있다.

[0020] 그리고, 상기 모음 개수 정보는 단어 내에서 모음의 개수가 하나 또는 둘 이상인지 여부에 따라 이진 분류될 수 있다.

[0021] 그리고, 상기 음차 표기 정보는 상기 분할된 음소에 대해 표기될 음차 표기에 대한 적어도 하나의 단위 표기 정보와; 상기 각 단위 표기 정보가 초성, 중성 및 종성 중 어디에 속하는지 여부에 대한 삼성 정보와; 현재 생성되는 음차 표기 정보의 대상이 되는 음소 이후의 음소 중 상기 (c) 단계의 수행을 스킵(skip)할 적어도 하나의 음소에 대한 스킵 음소 정보를 포함할 수 있다.

[0022] 그리고, 상기 표기 규칙 추출부는 상기 추출된 음소가 상기 복수의 음소 단위 음차 표기 규칙 중 어느 하나에 부합되는지 여부를 각 음소별로 상기 복수의 음소 단위 음차 표기 규칙에 대해 기 설정된 우선 순위에 따라 순차적으로 체크하고; 상기 복수의 음소 단위 음차 표기 규칙 중 상기 추출된 음소가 부합되는 최우선 순위의 어느 하나를 상기 분할된 음소에 해당하는 상기 음소 단위 음차 표기 규칙으로 추출할 수 있다.

[0023] 그리고, 상기 음차 표기 정보는 상기 분할된 음소의 음차 표기 방식에 대한 적어도 하나의 단위 표기 정보와; 상기 각 단위 표기 정보가 초성, 중성 및 종성 중 어디에 속하는지 여부에 대한 삼성 정보와; 현재 생성되는 음차 표기 정보의 대상이 되는 음소 이후의 음소 중 상기 표기 규칙 추출부가 상기 음소 단위 음차 표기 규칙의 추출을 스킵(skip)할 적어도 하나의 음소에 대한 스킵 음소 정보를 포함할 수 있다.

발명의 효과

[0024] 상기 구성에 따라 본 발명에 따르면, 음철법(Phonics)의 개념을 이용하여 음차 표기 규칙을 생성하여 이를 자동 음차 표기에 적용함으로써, 보다 정확하고 용례를 많이 수집하기 어려운 경우나 신조어 등의 처리에도 강인한 외래어의 자동 음차 표기 방법 및 장치가 제공된다.

도면의 간단한 설명

[0025] 도 1은 본 발명에 따른 외래어의 자동 음차 표기 장치의 구성을 도시한 도면이고,
 도 2 내지 도 9는 본 발명에 따른 외래어의 자동 음차 표기 장치에 적용된 음소 단위 음차 표기 규칙과 음차 표기 정보의 예를 도시한 도면이고,
 도 10은 본 발명에 따른 외래어의 자동 음차 표기 방법을 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0026] 이하에서는 첨부된 도면들을 참조하여 본 발명에 따른 실시예에 대해 상세히 설명한다.

[0027] 본 발명에 따른 외래어의 자동 음차 표기 장치는 음철법(Phonics)의 개념을 도입하여 음차 표기 규칙을 생성한다. 음철법(Phonics)은 영어권 국가에서 5-6세 정도의 말을 알아듣고 이해하지만 글자를 읽고 쓰지 못하는 아이들에게 문자를 가르치기 위해 만든 교수법으로, 한국에서도 음철법을 초등학교 영어 교육에 도입하여 한글을 읽을 수 있지만 영어를 읽지 못하는 아이들의 교육에 사용하고 있다.

[0028] 도 1은 본 발명에 따른 외래어의 자동 음차 표기 장치의 구성을 도시한 도면이다. 도 1을 참조하여 설명하면, 본 발명에 따른 자동 음차 표기 장치는 표기 규칙 데이터베이스(16), 표기 정보 데이터베이스(17), 음소 분할부(11), 표기 규칙 추출부(12), 표기 정보 생성부(13) 및 음차 표기 생성부(14)를 포함한다. 또한, 본 발명에 따른 자동 음차 표기 장치는 단어 입력부(10) 및 음차 표기 출력부(15)를 포함할 수 있다.

[0029] 단어 입력부(10)로는 본 발명에 따른 자동 음차 표기 장치에 의해 변환할 변환 대상 단어가 입력된다. 그리고, 음소 분할부(11)는 단어 입력부(10)를 통해 입력된 변환 대상 단어를 음소 단위로 분할한다. 여기서, 본 발명에 따른 자동 음차 표기 장치를 통해 음차 표기될 외래어는 영어인 것을 예로 하며, 음소 분할부(11)는 영어의 알파벳 단위를 음소 단위로 하여 변환 대상 단어를 분할한다.

- [0030] 표기 규칙 데이터베이스(16)에는 외래어를 구성하는 음소별로 복수개로 마련되는 음소 단위 음차 표기 규칙이 저장된다. 본 발명에 따른 음소 단위 음차 표기 규칙은 외래어를 구성하는 각 음소에 대한 타겟 음소 정보와, 각 음소의 주변 음소에 대한 주변 음소 정보에 기초하여 생성된다. 그리고, 표기 정보 데이터베이스(17)에는 표기 규칙 데이터베이스(16)에 저장된 복수의 음소 단위 음차 표기 규칙에 각각 대응하도록 마련된 음차 표기 정보가 저장된다.
- [0031] 이하에서는, 도 2 내지 도 9를 참조하여 본 발명에 따른 음소 단위 음차 표기 규칙과 음차 표기 정보에 대해 상세히 설명한다. 먼저, 본 발명에 따른 음소 단위 음차 표기 규칙의 생성에 사용되는 타겟 음소 정보는 타겟 음소(source), 타겟 자모 정보(jamo), 모음 개수 정보(moi), 모음 위치 정보를 포함할 수 있다.
- [0032] 타겟 음소(source)는 타겟 음소 정보의 대상이 되는 음소로, 상술한 바와 같이, 음차 표기될 외래어가 영어인 경우, 타겟 음소(source)는 각각의 알파벳에 해당하게 된다. 타겟 자모 정보(jamo)는 타겟 음소(source)가 자음인지 모음인지에 대한 정보를 갖게 된다. 도 2 내지 도 9에 도시된 바와 같이, 알파벳 a, e, l, o, u, w, y가 본 발명에 따른 음소 단위 음차 표기 규칙의 생성에서 모음으로 구분되고, 나머지 알파벳이 자음으로 구분되며, 모음을 'mo'로 표기하고, 자음을 'ja'로 표기하고 있다.
- [0033] 모음 개수 정보(moi)는 단어 내에서 모음의 개수에 대한 정보를 포함한다. 여기서, 모음 개수 정보(moi)에 사용되는 모음은 단어 내에서 소리가 나는 모음의 개수로, 본 발명에서는 모음 개수가 단어 내에서 모음의 개수가 하나 또는 둘 이상인지 여부에 따라 이진 분류되는 것을 예로 하고 있다. 예를 들어, 도 2 내지 도 9에 도시된 바와 같이, 단어 내에서 소리가 나는 모음의 개수가 하나일 경우는 '1'로 표기하고, 둘 이상일 경우에는 '2'로 표기할 수 있다.
- [0034] 모음 위치 정보는 현재 음차 표기하고자 하는 음소, 즉 타겟 음소(source)가 단어 내에서 몇 번째로 소리가 나는 모음인지에 대한 정보를 포함한다. 여기서, 모음 위치 정보는 타겟 음소(source)가 단어 내에서 첫 번째로 소리가 나는 모음인지 아닌지에 따라 이진 분류되는 것을 예로 하며, 도 2 내지 도 9에서는 모음 위치 정보로 타겟 음소(source)가 단어 내에서 첫 번째로 소리가 나는 모음인 경우 '1'로 표기하고, 두 번째 또는 그 이후의 모음인 경우 '2'로 표기하는 것을 예로 하고 있다.
- [0035] 한편, 주변 음소 정보는 콘텍스트 정보(context), 주변 자모 정보(cont-jamo) 및 단어 말단 정보(end)를 포함할 수 있다. 먼저, 콘텍스트 정보(context)는, 도 2 내지 도 9에 도시된 바와 같이, 타겟 주변 음소와 주변 음소 거리 정보를 포함할 수 있다.
- [0036] 타겟 주변 음소는 주변 음소 정보로 적용될 대상이 되는 음소로, 외래어 표기시 타겟 음소(source)의 주변에 표기되어 타겟 음소(source)의 발음을 가변시키는 음소로 설정될 수 있다. 보다 구체적으로 설명하면, 본 발명에 따른 자동 음차 표기 장치에 의해 음차 표기될 외래어가 영어인 경우, 변환 대상 단어인 영어 단어들에서 타겟 음소(source)인 알파벳의 주변에 기재되는 알파벳으로 타겟 음소(source)인 알파벳의 발음을 변화시키는 다른 알파벳이 타겟 주변 음소가 된다. 예를 들어, 도 2에서 타겟 음소(source)가 알파벳 'a'인 경우, 타겟 주변 음소로 'r' 'l' 'u' 등을 예로 하고 있다.
- [0037] 주변 음소 거리 정보는 타겟 음소(source)와 타겟 주변 음소 간의 거리에 대한 정보를 내포한다. 도 2에 도시된 타겟 음소(source) 'a'의 두 번째 음소 단위 음차 표기 규칙에 기재된 콘텍스트 정보(context)인 'l(en, +1), l(en, +2)'를 예로 하여 설명하면, 두 개의 'l'은 각각 타겟 주변 음소이고, 괄호 안의 '+1'과 '+2'가 주변 음소 거리 정보가 된다. 즉, 주변 음소 거리 정보는 타겟 음소(source)보다 앞에 나오는 음소일 경우에는 음수로 표기되고, 뒤에 나오는 음소인 경우에는 양수로 표기하며, 숫자가 타겟 음소(source)로부터 몇 번째 음소인지를 의미하게 된다.
- [0038] 여기서, 콘텍스트 정보(context)의 괄호 안에 기재된 'en'은 타겟 주변 음소의 언어 정보를 의미한다. 이는, 대부분의 경우 외래어의 문맥을 보게 되나 영어의 알파벳 'l'을 음차로 표기하면서 앞 음절이 모음으로 끝날 경우 반침으로 'ㄹ'을 삽입하여야 하므로 한국어 배열을 조건으로 사용할 경우가 발생하므로, 콘텍스트 정보(context)에 음소의 언어 정보를 포함시키게 된다. 본 발명에서는 고려하여야 할 음소가 외국어일 경우에는 'en'으로 표기하고, 고려하여야 할 음소가 한국어인 경우에는 'ko'로 표기하는 것을 예로 한다.
- [0039] 그리고, 도 2의 타겟 음소(source) 'a'의 첫 번째 음소 단위 음차 표기 규칙의 콘텍스트 정보(context)에 기재된 '/n(en, +2)'는 단어가 '~ar'로 끝나는 경우를 정의하고 있다. 즉, 'a'의 첫 번째 음소 단위 음차 표기 규칙은 단어가 '~ar'로 끝나는 경우의 단어에서 타겟 음소(source)가 'a'인 경우에 부합되는 규칙이 된다.

- [0040] 한편, 주변 음소 정보 중 주변 자모 정보(cont-jamo)는 타겟 주변 음소가 자음인지 모음인지 여부에 대한 정보를 포함한다. 이는 타겟 음소(source)가 타겟 주변 음소가 자음이거나 모음인 경우에 따라 그 음차 표기가 변할 수 있음을 반영한 경우로, 예를 들어 영어 단어의 경우 's' 뒤에 자음이 올 경우에는 '스'라고 음차 표기하는 예를 반영하고 있다.
- [0041] 여기서, 주변 자모 정보(cont-jamo)의 표기는 타겟 주변 음소가 자음인 경우에는 'ja'로 표기하고, 모음인 경우에는 'mo'로 표기하는 것을 예로 하고 있다. 그리고, 타겟 주변 음소가 타겟 음소(source)의 앞에 위치하는 음소인 경우에는 'ja' 또는 'mo'에 '-'를 부가하고, 타겟 음소(source)의 뒤에 위치하는 음소인 경우에는 '+'를 부가하는 것을 예로 하고 있다.
- [0042] 단어 말단 정보(end)는 타겟 음소(source)가 변환 대상 단어의 마지막 음소인지 여부에 대한 정보이다. 예를 들어, 영어 단어의 경우 's'가 단어의 마지막 음소일 경우 '스'라고 음차 표기하는 경우를 반영한 것으로, 마지막 음소가 자음일 경우 음차 표기에 '-'를 삽입 가능하게 하는 정보이다. 도 2 내지 도 9에서는 마지막 음소가 적용되는 음소 단위 음차 표기 규칙의 단어 말단 정보(end)로 'yes'를 표기하는 것을 예로 하고 있다.
- [0043] 상기와 같은 타겟 음소 정보와 주변 음소 정보에 기초하여 외래어를 구성하는 각 음소별로 복수의 음소 단위 음차 표기 규칙이 생성되며, 도 2 내지 도 9는 외래어가 영어인 경우로 음소 단위를 알파벳으로 하여, 각 알파벳에 대해 복수의 음소 단위 음차 표기 규칙을 생성한 예를 도시하고 있다.
- [0044] 여기서, 각 음소별로 생성된 복수의 음소 단위 음차 표기 규칙은 기 설정된 우선 순위를 갖게 된다. 도 2 내지 도 9에서 각 음소별로 우선 순위가 높은 순서로 음소 단위 음차 표기 규칙을 표로 기재한 것이다.
- [0045] 또한, 각 음소 단위 음차 표기 규칙은 고려 대상 정보와 비고려 대상 정보로 구분되는데, 모음 개수 정보(moi), 모음 위치 정보, 콘텍스트 정보(context), 주변 자모 정보(cont-jamo) 및 단어 말단 정보(end)가 음소별, 그리고 음소 단위 음차 표기 규칙별로 고려 대상 정보와 비고려 대상 정보에 선택적으로 포함된다. 여기서, 모음 개수 정보(moi), 모음 위치 정보, 콘텍스트 정보(context), 주변 자모 정보(cont-jamo) 및 단어 말단 정보(end) 중 특정 음소 단위 음차 표기 규칙에서 고려 대상 정보로 적용된 정보에는 상술한 바와 같은 표기 방식이 적용되며, 비고려 대상 정보로 적용되는 정보는 '0'으로 기재된다.
- [0046] 예를 들어, 타겟 음소(source)의 앞이나 뒤에 오는 음소가 단지 모음인지 자음인지 여부만 고려되어지는 경우에는 콘텍스트 정보(context)가 음차 표기에 영향을 미치지 않는 비고려 대상 정보로 적용되어 '0'으로 표기될 수 있다. 또한, 단어 내에서의 모음 개수나 모음의 위치가 음차 표기에 영향을 미치지 않는 경우, 해당 음소 단위 음차 표기 규칙에서 모음 개수 정보(moi)나 모음 위치 정보가 '0'으로 표기된다.
- [0047] 다시, 도 1을 참조하여 설명하면, 상기와 같이 방법으로 구축된 음소 단위 음차 표기 규칙에 표기 규칙 데이터베이스(16)에 등록된 상태에서, 표기 규칙 추출부(12)는 표기 규칙 데이터베이스(16)로부터 음소 분할부(11)에 의해 분할된 각 음소에 대응하는 음소 단위 음차 표기 규칙을 추출한다.
- [0048] 여기서, 본 발명에 따른 표기 규칙 추출부(12)는, 도 1에 도시된 바와 같이, 단어 분석부(12a)와 표기 규칙 검색부(12b)를 포함할 수 있다. 단어 분석부(12a)는 음소 분할부(11)에 의해 분할된 변환 대상 단어의 음소들을 분석하여, 해당 변환 대상 단어의 각 음소, 모음의 개수, 해당 모음의 위치, 주변 음소에 대한 정보 등을 분석한다.
- [0049] 그리고, 표기 규칙 검색부(12b)는 단어 분석부(12a)에 의해 분석된 변환 대상 단어의 각 음소들에 대한 정보를 이용하여, 표기 규칙 데이터베이스(16)로부터 해당 음소에 대응하는 음소 단위 음차 표기 규칙을 검색하고, 검색된 음소 단위 음차 표기 규칙을 해당 음소에 대한 음소 단위 음차 표기 규칙으로 추출하게 된다.
- [0050] 여기서, 표기 규칙 추출부(12)는 음차 표기 대상 음소가 해당 음소에 대해 등록된 복수의 음소 단위 음차 표기 규칙 중 어디에 부합되는지 여부를 복수의 음소 단위 음차 표기 규칙에 대해 설정된 상술한 우선 순위에 따라 순차적으로 체크하고, 부합되는 것으로 판단된 최우선 순위의 음소 단위 음차 표기 규칙을 해당 음소에 대한 음소 단위 음차 표기 규칙으로 추출하게 된다.
- [0051] 예를 들어, 음차 표기 대상 음소가 'a'인 경우, 'a'를 포함하는 변환 대상 단어가, 도 2에 도시된 'a'에 대한 13개의 음소 단위 음차 표기 규칙 중 어느 규칙에 부합되는지 여부를 우선 순위, 즉 도 2의 기재 순서에 따라 체크하고, 부합되는 규칙이 나타나는 경우 해당 규칙을 변환 대상 단어의 음소인 'a'에 대한 음소 단위 음차 표기 규칙으로 추출하게 된다.
- [0052] 영어 단어 'end'를 예로 하면, 음소 분할부(11)에 의해 분할된 음소는 'e', 'n', 'd' 3개가 되며, 첫 번째 음소

인 'e'에 대한 음소 단위 음차 표기 규칙을 추출할 때, 도 2에 도시된 'e'에 대한 첫 번째 음소 단위 음차 표기 규칙이 체크된다.

- [0053] 이 때, 도 2에 도시된 'e'에 대한 첫 번째 음소 단위 음차 표기 규칙에서는 모음 개수 정보(moi)와 모음 위치 정보가 비교 대상 정보에 해당되어 변환 대상 단어 'end'에서 모음 개수나 모음 위치는 고려되지 않게 된다. 그리고, 콘텍스트 정보(context)와 관련하여, 변환 대상 단어 'end'의 타겟 음소(source)인 'e'의 타겟 주변 음소에는 'r'이 존재하지 않으므로, 'e'에 대한 첫 번째 음소 단위 음차 표기 규칙은 변환 대상 단어 'end' 중 타겟 음소(source) 'e'에 부합하는 음소 단위 음차 표기 규칙이 아닌 것으로 판단된다.
- [0054] 상기와 같은 과정을 통해, 순차적인 체크 과정 중, 도 2에 도시된 'e'에 대한 11 번째 음소 단위 음차 표기 규칙이 변환 대상 단어 'end' 중 타겟 음소(source) 'e'에 부합하는 음소 단위 음차 표기 규칙으로 추출된다. 즉, 변환 대상 단어 'end'에는 모음이 하나이고, 타겟 음소(source) 'e'가 첫 번째 모음이며, 바로 뒤 음소가 자음인 'n'이므로 도 2에 도시된 'e'에 대한 11 번째 음소 단위 음차 표기 규칙에 부합된다.
- [0055] 상기와 같은 과정을 통해, 표기 규칙 추출부(12)가 변환 대상 단어의 각 음소에 대응하는 음소 단위 음차 표기 규칙을 추출하면, 표기 정보 생성부(13)는 표기 규칙 추출부(12)에 의해 추출된 음소 단위 음차 표기 규칙에 대응하는 음차 표기 정보를 표기 정보 데이터베이스(17)로부터 추출하여 각 음소에 대한 음차 표기 정보를 생성한다.
- [0056] 여기서, 표기 정보 데이터베이스(17)에 저장되는 음차 표기 정보는, 도 2 내지 도 9에 도시된 바와 같이, 표기 규칙 데이터베이스(16)에 저장된 각 음소 단위 음차 표기 규칙에 매칭되도록 마련된다. 도 2 내지 도 9를 참조하여 설명하면, 본 발명에 따른 음차 표기 정보는 적어도 하나의 단위 표기 정보(target), 삼성 정보 및 스킵 음소 정보(eni)를 포함할 수 있다.
- [0057] 음차 표기 정보의 단위 표기 정보(target)는 타겟 음소(source)의 음차 표기 방식을 정의한다. 도 2에 도시된 'a'의 8 번째 음소 단위 음차 표기 규칙에 대한 음차 표기 정보를 예로 하여 설명하면, 단위 표기 정보(target)는 'ㄱ', 'ㅇ', 'ㅣ'를 포함하게 된다. 즉, 단위 표기 정보(target)는 타겟 음소의 실질적인 음차 표기를 정의한다.
- [0058] 삼성 정보는 단위 표기 정보(target), 즉, 실질적인 음차 표기가 한국어의 초성, 중성 및 종성 중 어디에 해당하는지 여부를 정의하며, 초성은 '1', 중성은 '2', 종성은 '3'으로 표기한다. 도 2 내지 도 9에서는 괄호 안에 첫 번째로 삼성 정보를 표기하는 것을 예로 하고 있으며, 도 2에 도시된 'a'의 8 번째 음소 단위 음차 표기 규칙에 대한 음차 표기 정보에서는 단위 표기 정보(target) 'ㄱ'가 중성, 'ㅇ'가 초성, 'ㅣ'가 중성임을 정의하고 있다. 따라서, 'a'가 음차 표기로 '에이'로 표기됨을 정의하게 된다. 여기서, 괄호 안의 삼성 정보 다음에 기재되는 숫자는 해당 단위 표기 정보(target), 즉 실질적인 음차 표기의 위치, 즉 단위 표기 정보(target)의 순서를 정의하고 있다.
- [0059] 한편, 스킵 음소 정보(eni)는 변환 대상 단어의 음소들 중 현재 생성되는 음차 표기 정보의 대상이 되는 음소 이후의 음소 중 표기 규칙 추출부(12)가 음소 단위 표기 규칙의 추출을 스킵(skip)할 적어도 하나의 음소를 정의한다.
- [0060] 예를 들어, 도 2에 도시된 'a'의 8 번째 음소 단위 음차 표기 규칙에서는 변환 대상 단어 중 연속된 음소 '~ay~'와 관련하여, 타겟 음소(source)가 'a'이고, 타겟 주변 음소가 'y'일 때, 음차 표기는 상술한 바와 같이 '에이'가 된다. 이 때, 음차 표기 '에이'는 변환 대상 단어의 연속된 음소 '~ay~'에 대한 음차 표기이므로 'y'에 대한 음소 단위 음차 표기 규칙의 추출이 불필요하게 된다.
- [0061] 따라서, 타겟 음소(source)가 'a'에 대한 음차 표기 정보에 스킵 음소 정보(eni)를 포함시킴으로써, 타겟 음소(source) 이후의 음소 중 음소 단위 표기 규칙의 추출을 스킵(skip)할 음소를 정의하게 된다. 여기서, 도 2 내지 도 9에서는 스킵 음소 정보(eni)를 스킵(skip)할 음소의 개수로 양수로 표기하는 것을 일 예로 하고 있다.
- [0062] 그리고, 도 2 내지 도 8에 도시된 음소 표기 정보의 'koi'는 다음 연산에서 건너뛸 한국어, 즉 음차 표기를 정의한 것으로, 하나의 타겟 음소(source)에 여러개의 한국어 음소가 음차 표기로 결정될 경우, 해당 한국어 음소 개수만큼 건너뛸 이후의 배열에 다음 타겟 음소(source)에 대한 음차 표기가 기록된다. 현재 타겟 음소(source)와 다음 타겟 음소(source) 사이에는 현재 타겟 음소(source)에 대한 음차 표기가 채워짐은 물론이다.
- [0063] 상기와 같이 표기 정보 생성부(13)가 변환 대상 단어의 각 음소에 대한 음차 표기 정보를 생성하게 되면, 음차 표기 생성부(14)가 표기 정보 생성부(13)에 의해 생성된 음차 표기 정보에 기초하여 변환 대상 단어에 대한 음

차 표기를 생성한다. 그리고, 음차 표기 생성부(14)에 의해 생성된 음차 표기는 음차 표기 출력부(15)를 통해 출력된다.

[0064] 상기와 같은 구성을 통해, 영어 단어 'youngstown'을 본 발명에 따른 자동 음차 표기 장치에 적용하는 경우, [표 1]과 같은 결과를 얻게 되며, 음차 표기는 '영스타운'이 된다.

[0065] [표 1]

source	jamo	moc	moi	context	cont-jamo	end	target	koi	eni
y	mo	0	0	o(en,+1),u(en,+2) n(en,+3),g(en,+4)	0	0	ㅃ (2,0)	0	+2
o	skip								
u	skip								
n	ja	0	0	g(en,+4)	0	0	ㅇ (3,0)	0	+1
g	skip								
s	ja	0	0	0	0	0	ㅅ (1,0), ㅡ (2,+1)	+1	0
t	ja	0	0	0	ㅌmo	0	ㅌ (1,0)	0	0
o	mo	0	0	w(en,+1)	0	0	ㅍ (2,0), ㅇ (1,+1) ㅌ (2,+2)	+2	+1
w	skip								
n	ja	0	0	0	ㄴmo	0	ㄴ (3,0)	0	0

[0066]

이하에서는, 도 10을 참조하여 본 발명에 따른 외래어의 자동 음차 표기 방법에 대해 설명한다.

[0067]

[0068] 먼저, 상술한 바와 같이 복수의 음소 단위 음차 표기 규칙과, 이에 대응하는 음차 표기 정보가 각각 표기 규칙 데이터베이스(16) 및 표기 정보 데이터베이스(17)에 구축된 상태(S20)에서, 변환 대상 단어가 본 발명에 따른 자동 음차 표기 장치에 입력된다(S21).

[0069]

그런 다음, 입력된 변환 대상 단어가 음소 분할부(11)에 의해 음소 단위로 분할된다(S22). 그리고, 분할된 첫 번째 음소가 모음인지 여부를 체크하여(S23), 모음인 경우 해당 모음이 변환 대상 단어의 첫 음소인지 여부를 판단한다(S24).

[0070]

여기서, S24 단계에서 해당 모음이 변환 대상 단어의 첫 음소인 것으로 판단되는 경우, 초성에 'ㅇ'을 삽입한 후, 해당 음소에 대하여 음소 단위 음차 표기 규칙을 추출한다(S26). 그리고, 추출된 음소 단위 음차 표기 규칙에 대응하는 음차 표기 정보가 생성된다(S27). 여기서, 음소 단위 음차 표기 규칙의 추출과 음차 표기 정보의 생성 과정은 상술한 바와 같은 바, 그 상세한 설명은 생략한다.

[0071]

상기와 같은 과정을 통해 하나의 음소에 대하여 음차 표기 정보가 생성되면, S28 단계에서 해당 음소가 단어의 마지막 음소인지 여부를 체크하고(S28), 마지막 음소가 아닌 경우에는 다음 음소에 대하여 S23 단계 내지 S27 단계를 수행하여 음차 표기 정보를 생성한다.

[0072]

여기서, S23 단계 및 S24 단계에서 각각 해당 음소가 모음이 아닌 것으로 판단되거나, 첫 음소가 아닌 것으로 판단된 경우에는 해당 음소의 스킵(skip) 여부를 판단한다(S30). 여기서, S30 단계에서의 스킵(skip) 여부는 상술한 음차 표기 정보의 스킵 음소 정보(eni)에 기초하여 결정되며, 그 상세한 설명은 상술한 바와 같다.

[0073]

상기 과정을 통해 변환 대상 단어의 모든 음소에 대하여 음차 표기 정보의 생성이 완료되면, 음차 표기 생성부(14)가 변환 대상 단어에 대한 음차 표기를 생성하게 된다(S29).

[0074]

이하에서는 본 발명에 따른 외래어의 자동 음차 표기 방법에 따른 음차 표기의 정확성을 검증하기 위한 과정과 그 결과에 대해 설명한다.

[0075]

본 발명에 따른 외래어의 자동 음차 표기 방법의 정확성 검증을 위해 국립 국어원에 공개되어 있는 외래어 표기 용례 자료 중 로마자 인명과 지명을 평가 집합으로 사용한다. 다만, "Prince Andrew, 앤드루 왕자"와 같이 여

러 단어로 되어 있는 경우 띄어쓰기를 기준으로 분리를 하고 "Prince"를 "왕자"로 표기한 것은 음차 표기가 아닌 번역이므로 이런 경우는 수동으로 찾아서 평가 집합에서 제외하였다. 이러한 과정을 통해 최종 평가 집단으로 총 15,651개의 단어가 선정되었다.

[0076] 평가의 척도로는 NEWS 2009 Machine Transliteration Shared Task(<http://www.acl-ijcnlp-2009.org/workshops/NEWS2009/pages/sharedtask.html> 참조)에서 사용한 평가 척도 중 평균 F-measure를 사용하는 것을 예로 하였으며, 한글은 초성, 중성, 종성을 모아쓰기를 하므로 이를 분해하여 음소 단위로 F-measure를 측정하였다.

[0077] 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 산출되며, 정답 단어를 r , 시스템의 결과를 c 라고 하고 편집 거리를 $ED(c, r)$ 이라고 했을 때 최장 공통 음소열은 [수학식 1]과 같이 표현된다.

[0078] [수학식 1]

[0079]
$$LCS(c, r) = 1/2(\text{length}(c) + \text{length}(r) - ED(c, r))$$

[0080] 여기서, $LCS(c, r)$ 는 최장 공통 음소열이고, $\text{length}(c)$ 는 정답 단어의 길이이고, $\text{length}(r)$ 은 시스템 결과의 길이이다.

[0081] 그리고, 최장 공통 음소열을 이용하여 단어 i 의 정확률(P_i)과 재현율(R_i)을 측정하면, [수학식 2]와 [수학식 3]과 같이 표현된다.

[0082] [수학식 2]

[0083]
$$P_i = \frac{LCS(c_i, r_i)}{\text{length}(c_i)}$$

[0084] [수학식 3]

[0085]
$$R_i = \frac{LCS(c_i, r_i)}{\text{length}(r_i)}$$

[0086] 따라서, 단어 i 의 음소 단위 F-measure는 [수학식 4]를 통해 측정 가능하게 된다.

[0087] [수학식 4]

[0088]
$$F_i = 2 \frac{R_i \times P_i}{R_i + P_i}$$

[0089] 여기서, 음소 단위 F-measure의 값이 1에 가까울수록 시스템의 결과가 정답과 비슷함을 의미하게 되고, F-measure 값이 1이 되면 정확히 일치함을 의미하게 된다.

[0090] 본 발명에 따른 외래어의 자동 음차 표기 방법을 통해 상술한 평가 집단의 15,651개의 단어를 음차 표기해본 결과, F-measure가 0.766로 측정되었다. 이 중 F-measure가 0.5 미만은 999개였으며, 3,019개 단어는 정확한 음차 표기가 되었음을 확인하였다. 또한, 오류를 분석한 결과 [표 2]에 기재된 바와 같이 영어 이외의 언어권의 고유 명사가 F-measure를 기준으로 하위권에 있음을 확인할 수 있었다.

[0091] [표 2]

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

[0092]

[0093] 이는 본 발명에 따른 외래어의 자동 음차 표기 방법에 사용되는 음소 단위 음차 표기 규칙을 영어를 대상으로 한 음절법을 기초로 구축한 데 기인한 결과로, 평가 집단으로 사용된 외래어 표기 용례 자료에는 영어 이외의 로마자 표기 언어도 포함되어 있기 때문이다. 이와 같은 단어의 경우에는 해당 언어에 적합한 음소 단위 음차 표기 규칙을 추가하는 경우 정확한 음차 표기가 가능하게 된다.

[0094] 이상에서 본 발명의 바람직한 실시예에 대하여 상세하게 설명하였지만 본 발명의 권리범위는 이에 한정되는 것은 아니고 다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한 당업자의 여러 변형 및 개량 형태 또한 본 발명의 권리범위에 속하는 것이다.

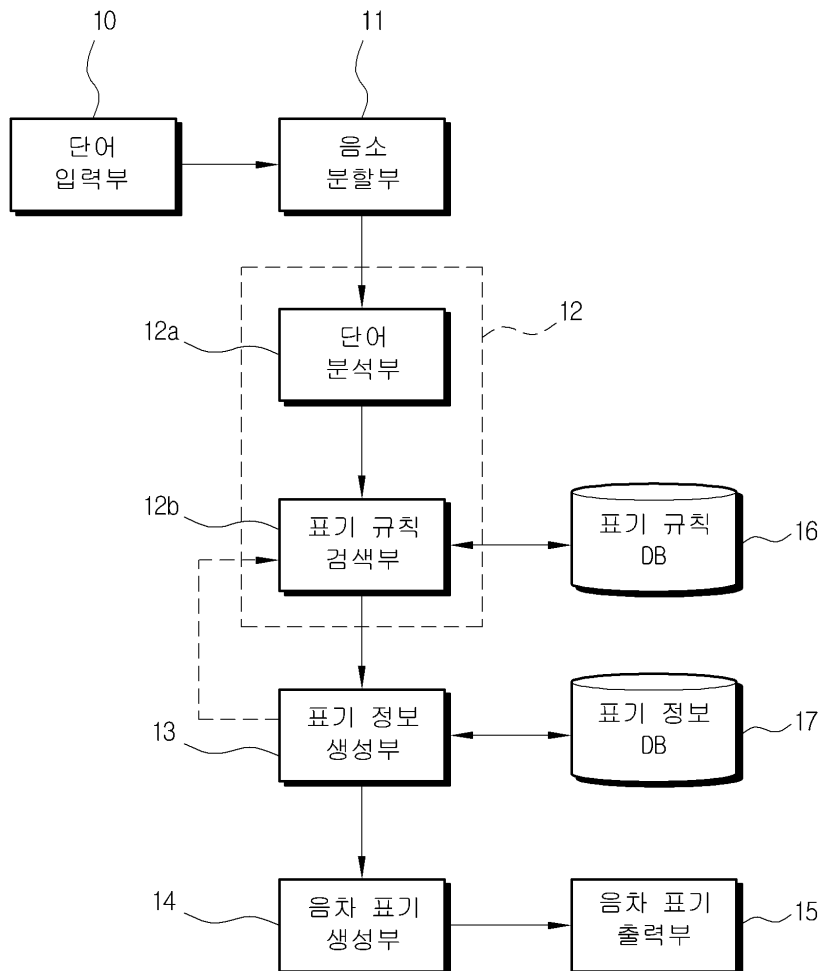
부호의 설명

[0095]

10 : 단어 입력부	11 : 음소 분할부
12 : 표기 규칙 추출부	13 : 표기 정보 생성부
14 : 음차 표기 생성부	15 : 음차 표기 출력부
16 : 표기 규칙 데이터베이스	17 : 표기 정보 데이터베이스

도면

도면1



도면2

source	jamo	moc	moi	context	cont - jamo	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	ㅑ(2,0)	0	0
	mo	0	0	l(en,+1),t(en,+2)	0	0	ㅓ(2,0)	0	0
	mo	0	0	l(en,+1),k(en,+2)	0	0	ㅕ(2,0)	0	0
	mo	0	0	l(en,+1)	0	0	ㅗ(2,0)	0	0
	mo	0	0	u(en,+1)	0	0	ㅛ(2,0)	0	+1
	mo	0	0	w(en,+1)	0	0	ㅜ(2,0)	0	+1
	mo	0	0	y(en,+1)	0	0	ㅟ(2,0),ㅇ(1,+1), ㅣ(2,+2)	+2	+1
	mo	0	0	h(en,+1),/n(en,+2)	0	0	ㅑ(2,0)	0	+1
	mo	1	1	0	+ja	no	ㅓ(2,0)	0	0
	mo	1	1	0	0	0	ㅟ(2,0),ㅇ(1,+1), ㅣ(2,+2)	+2	0
	mo	2	0	0	+ja	0	ㅏ(2,0)	0	0
	mo	2	0	0	0	0	ㅟ(2,0),ㅇ(1,+1), ㅣ(2,+2)	+2	+1
e	mo	0	0	r(en,+1),/n(en,+2)	0	0	ㅑ(2,0)	0	+1
	mo	0	0	r(en,+1)	0	0	ㅑ(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	ㅟ(2,0),ㅇ(1,+1), ㅣ(2,+2)	+2	+3
	mo	0	0	i(en,+1)	0	0	ㅟ(2,0),ㅇ(1,+1), ㅣ(2,+2)	+2	+1
	mo	0	0	w(en,+1)	0	0	ㅟ(2,0)	0	+1
	mo	0	0	o(en,+1)	0	0	ㅑ(2,0)	0	+1
	mo	0	0	u(en,+1)	0	0	ㅟ(2,0)	0	+1
	mo	0	0	h(en,-1),t(en,-2)	0	0	ㅑ(2,0)	0	0
	mo	1	1	a(en,+1),r(en,+2)	0	0	ㅣ(2,0),ㅇ(1,+1), ㅑ(2,+2)	+2	+2
	mo	1	1	a(en,+1)	0	0	ㅣ(2,0)	0	+1
	mo	1	1	0	+ja	0	ㅓ(2,0)	0	0
	mo	1	1	0	0	0	ㅣ(2,0)	0	0
	mo	2	0	a(en,+1),r(en,+2)	0	0	ㅣ(2,0),ㅇ(1,+1), ㅑ(2,+2)	+2	+1
	mo	2	0	e(en,+1),r(en,+2)	0	0	ㅣ(2,0),ㅇ(1,+1), ㅑ(2,+2)	+2	+1

도면3

source	jamo	moc	moi	context	cont - jamo	end	target	koi	eni
e	mo	2	0	y(en,+1),r(en,+2)	0	0	l(2,0),o(1,+1), j(2,+2)	+2	+1
	mo	2	0	a(en,+1)	0	0	l(2,0)	0	+1
	mo	2	0	e(en,+1)	0	0	l(2,0)	0	+1
	mo	2	0	y(en,+1)	0	0	l(2,0)	0	+1
	mo	2	0	0	0	0	l(2,0)	0	0
i	mo	1	1	r(en,+1)	+mo	0	l(2,0),o(1,+1), j(2,+2)	+2	+1
	mo	2	0	r(en,+1)	+ja	0	j(2,0)	0	+1
	mo	0	0	n(en,+1),d(en,+2)	0	0	l(2,0),o(1,+1), j(2,+2)	+2	+1
	mo	0	0	l(en,+1),d(en,+2)	0	0	l(2,0),o(1,+1), j(2,+2)	+2	+1
	mo	0	0	e(en,+1)	0	0	l(2,0)	0	+1
	mo	0	0	i(en,+1)	0	0	l(2,0)	0	+1
	mo	0	0	a(en,+1),h(en,+2), /n(en,+3)	0	0	l(2,0),o(2,+1), j(2,+2)	+2	+2
	mo	0	0	a(en,+1)	0	0	l(2,0),o(2,+1), j(2,+2)	+2	+1
	mo	0	0	o(en,+1)	0	0	l(2,0),o(2,+1), j(2,+2)	+2	+1
	mo	0	0	u(en,+1)	0	0	l(2,0),o(2,+1), T(2,+2)	+2	+1
	mo	1	1	0	+ja	0	l(2,0)	0	0
	mo	1	1	0	0	0	l(2,0),o(1,+1), j(2,+2)	+2	0
	mo	2	0	0	0	0	l(2,0)	0	0
o	mo	0	0	o(en,+1)	0	0	T(2,0)	0	+1
	mo	0	0	u(en,+1)	0	0	l(2,0),o(1,+1), T(2,+2)	+2	+1
	mo	0	0	w(en,+1)	0	0	l(2,0),o(1,+1), T(2,+2)	+2	+1
	mo	0	0	y(en,+1)	0	0	l(2,0),o(1,+1), j(2,+2)	+2	+1

도면4

source	jamo	moc	moi	context	cont - jamo	end	target	koi	eni
o	mo	0	0	i(en,+1)	0	0	$\perp(2,0), \circ(1,+1),$ $\mid(2,+2)$	+2	+1
	mo	0	0	a(en,+1)	0	0	$\perp(2,0), \circ(1,+1),$ $\vdash(2,+2)$	+2	+1
	mo	0	0	e(en,+1)	0	0	$\perp(2,0), \circ(1,+1),$ $\dashv(2,+2)$	+2	+1
	mo	0	0	0	0	0	$\perp(2,0)$	0	0
u	mo	0	0	r(en,+1)	+mo	0	$\dashv(2,0), \equiv(1,+1)$	+1	+1
	mo	0	0	r(en,+1)	0	0	$\dashv(2,0)$	0	+1
	mo	0	0	y(en,+1)	0	0	$\vdash(2,0), \circ(1,+1),$ $\mid(2,+2)$	+2	+1
	mo	1	1	0	+ja	0	$\top(2,0)$	0	0
	mo	1	1	0	0	0	$\top(2,0)$	0	0
	mo	2	0	0	0	0	$\top(2,0)$	0	0
	mo	0	0	0	0	0	$\top(2,0)$	0	0
	mo	0	0	0	0	0	$\top(2,0)$	0	0
w	mo	0	0	a(en,+1),r(en,+2)	0	0	$\dashv(2,0)$	0	+2
	mo	0	0	a(en,+1)	0	0	$\vdash(2,0)$	0	+1
	mo	0	0	e(en,+1),a(en,+2)	0	0	$\dashv(2,0)$	0	+2
	mo	0	0	e(en,+1)	0	0	$\dashv(2,0)$	0	+1
	mo	0	0	i(en,+1)	0	0	$\dashv(2,0)$	0	+1
	mo	0	0	o(en,+1)	0	0	$\perp(2,0)$	0	+1
	mo	0	0	u(en,+1)	0	0	$\top(2,0)$	0	+1
	mo	0	0	h(en,+1),e(en,+2)	0	0	$\dashv(2,0)$	0	+2
	mo	0	0	h(en,+1),a(en,+2)	0	0	$\dashv(2,0)$	0	+2
	mo	0	0	h(en,+1),i(en,+2)	0	0	$\dashv(2,0)$	0	+2
	mo	0	0	h(en,+1),o(en,+2)	0	0	$\equiv(1,-1), \top(2,0)$	+1	+2
	mo	0	0	h(en,+1),y(en,+2)	0	0	$\dashv(2,0), \circ(1,+1),$ $\mid(2,+2)$	+2	+2
	mo	0	0	0	0	0	$\top(2,0)$	0	0
	mo	0	0	0	0	0	$\top(2,0)$	0	0
y	mo	0	0	a(en,+1),e(en,+2)	0	0	$\dashv(2,0)$	0	+2
	mo	0	0	a(en,+1)	0	0	$\vdash(2,0)$	0	+1
	mo	0	0	e(en,+1),o(en,+2)	0	0	$\dashv(2,0)$	0	+1
	mo	0	0	e(en,+1)	0	0	$\dashv(2,0)$	0	+1
	mo	0	0	i(en,+1)	0	0	$\mid(2,0)$	0	+1
	mo	0	0	o(en,+1),u(en,+2), n(en,+3),g(en,+4)	0	0	$\dashv(2,0)$	0	+2
	mo	0	0	o(en,+1),u(en,+2)	0	0	$\top(2,0)$	0	+2

도면5

source	jamo	moc	moi	context	cont - jamo	end	target	koi	eni
y	mo	0	0	o(en,+1)	0	0	ㅛ(2,0)	0	+1
	mo	0	0	u(en,+1)	0	0	ㅠ(2,0)	0	+1
	mo	2	2	0	0	0	ㅣ(2,0)	0	0
	mo	1	1	0	0	0	ㅛ(2,0), ㅇ(1,+1), ㅣ(2,+2)	+2	0
	mo	0	0	0	0	0	ㅣ(2,0)	0	0
b	ja	0	0	0	-ja	yes	ㅛ(1,0), -(2,+1)	+1	0
	ja	0	0	0	-ja,+j a	0	ㅛ(1,0), -(2,+1)	+1	0
	ja	0	0	0	-ja,+m o	0	ㅛ(1,0)	0	0
	ja	0	0	0	+mo	0	ㅛ(1,0)	0	0
	ja	0	0	0	-mo	0	ㅛ(3,0)	0	0
	ja	0	0	0	0	0	ㅛ(1,0), -(2,+1)	+1	0
c	ja	0	0	e(en,+1)	0	0	ㅓ(1,0)	0	0
	ja	0	0	i(en,+1)	0	0	ㅕ(1,0)	0	0
	ja	0	0	y(en,+1)	0	0	ㅓ(1,0)	0	0
	ja	0	0	a(en,+1)	0	0	ㅕ(1,0)	0	0
	ja	0	0	o(en,+1)	0	0	ㅕ(1,0)	0	0
	ja	0	0	u(en,+1)	0	0	ㅕ(1,0)	0	0
	ja	0	0	h(en,+1), e(en,+2)	0	0	ㅓ(1,0)	0	+1
	ja	0	0	h(en,+1), i(en,+2)	0	0	ㅓ(1,0)	0	+1
	ja	0	0	h(en,+1), y(en,+2)	0	0	ㅓ(1,0)	0	+1
	ja	0	0	h(en,+1), o(en,+2)	0	0	ㅕ(1,0)	0	+1
	ja	0	0	h(en,+1), Wn(en,+2)	0	0	ㅕ(1,0), -(2,+1)	+1	+1
	ja	0	0	h(en,+1), a(en,+2)	0	0	ㅕ(1,0)	0	+1
	ja	0	0	h(en,+1), w(en,+2), s(en,-1)	0	0	ㅓ(1,0)	0	+1
	ja	0	0	h(en,+1)	+ja	0	ㅕ(1,0), -(2,+1)	+1	+1
	ja	0	0	k(en,+1), Wn(en,+2)	-mo	0	ㅕ(3,0)	0	+1
	ja	0	0	k(en,+1), Wn(en,+2)	0	0	ㅕ(1,0)	0	+1
	ja	0	0	k(en,+1)	+ja,-m o	0	ㅕ(3,0)	0	+1
	ja	0	0	k(en,+1)	+ja	0	ㅕ(1,0)	0	+1
	ja	0	0	k(en,+1)	0	0	ㅕ(1,0)	0	+1
	ja	0	0	c(en,+1), Wn(en,+2)	0	0	ㅕ(1,0), -(2,+1)	+1	+1
	ja	0	0	c(en,+1)	+ja	0	ㅕ(1,0), -(2,+1)	+1	+1

도면6

source	jamo	moc	moi	context	cont - jamo	end	target	koi	eni
c	ja	0	0	0	$\neg$ mo	yes	$\neg(3,0)$	0	0
	ja	0	0	0	0	yes	$\neg(1,0), \neg(2,+1)$	+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	$\neg(1,0), \neg(2,+1)$	+1	0
	ja	0	0	0	+ja	0	$\neg(1,0), \neg(2,+1)$	+1	0
	ja	0	0	0	0	0	$\neg(1,0)$	0	0
	ja	0	0	0	0	0	$\neg(1,0)$	0	0
f	ja	0	0	$f(en,+1)$	+ja	0	$\neg(1,0), \neg(2,+1)$	+1	+1
	ja	0	0	$f(en,+1)$	0	yes	$\neg(1,0), \neg(2,+1)$	+1	+1
	ja	0	0	0	+ja	0	$\neg(1,0), \neg(2,+1)$	+1	0
	ja	0	0	0	0	yes	$\neg(1,0), \neg(2,+1)$	+1	0
	ja	0	0	0	0	0	$\neg(1,0)$	0	0
g	ja	0	0	$h(en,+1)$	+ja	0	$\neg(1,0), \neg(2,+1)$	+1	+1
	ja	0	0	$h(en,+1)$	0	yes	$\neg(1,0), \neg(2,+1)$	+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	$\neg(1,0), \neg(2,+1)$	+1	+1
	ja	0	0	$r(en,+1)$	0	yes	$\neg(1,0), \neg(2,+1)$	+1	+1
	ja	0	0	$r(en,+1)$	$\neg$ mo	0	$\neg(1,0), \neg(2,+1)$	+1	0
	ja	0	0	$n(en,+1)$	0	0	$\neg(1,0), \neg(2,+1), \neg(3,+2)$	+2	+1
	ja	0	0	$l(en,+1)$	$\neg$ mo	0	$\neg(3,0)$	0	+1
	ja	0	0	0	+ja	0	$\neg(1,0), \neg(2,+1)$	+1	0
h	ja	0	0	0	0	yes	$\neg(1,0), \neg(2,+1)$	+1	0
	ja	0	0	0	0	0	$\neg(1,0)$	0	0
	ja	0	0	0	0	0	$\neg(1,0)$	0	0
j	ja	0	0	0	+ja	0	$\neg(1,0), \neg(2,+1)$	+1	0
	ja	0	0	0	0	yes	$\neg(1,0), \neg(2,+1)$	+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	$\neg(1,0), \neg(2,+1)$	+1	0
	ja	0	0	0	0	yes	$\neg(1,0), \neg(2,+1)$	+1	0

도면7

source	jamo	moc	moi	context	cont - jamo	end	target	koi	eni
k	ja	0	0	0	0	0	$\neg(1,0)$	0	0
l	ja	0	0	$l(en,+1)$	$\neg mo, +m$ o	0	$\exists(3,0)\exists(1,+1)$	+1	+1
	ja	0	0	$m(en,+1)$	$\neg mo, +m$ o	0	$\exists(3,0)\exists(1,+1)$	+1	0
	ja	0	0	0	$\neg mo, +m$ o	0	$\exists(3,0)\exists(1,+1)$	+1	0
	ja	0	0	0	+ja	0	$\exists(1,0), \neg(2,+1)$	0	0
	ja	0	0	0	0	yes	$\exists(1,0), \neg(2,+1)$	0	0
	ja	0	0	0	0	0	$\exists(1,0)$	0	0
m	ja	0	0	$m(en,+1)$	$\neg mo, +m$ o	no	$\Box(3,0), \Box(1,+1)$	+1	+1
	ja	0	0	$m(en,+1)$	$\neg mo$	0	$\Box(3,0)$	0	+1
	ja	0	0	$m(en,+1)$	0	0	$\Box(1,0)$	0	+1
	ja	0	0	0	+ja	0	$\Box(1,0), \neg(2,+1)$	+1	0
	ja	0	0	0	0	yes	$\Box(1,0), \neg(2,+1)$	+1	0
	ja	0	0	0	0	0	$\Box(1,0)$	0	0
n	ja	0	0	$g(en,+1), e(en,+2)$	0	0	$\circ(3,0), \times(1,+1)$	+1	+1
	ja	0	0	$g(en,+1), i(en,+2)$	0	0	$\circ(3,0), \times(1,+1)$	+1	+1
	ja	0	0	$g(en,+1), y(en,+2)$	0	0	$\circ(3,0), \times(1,+1)$	+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	$\circ(3,0), \neg(1,+1), \neg(2,+2)$	+2	+1
	ja	0	0	0	$\neg mo$	0	$\Box(3,0)$	0	0
	ja	0	0	0	+mo	0	$\Box(1,0)$	0	0
p	ja	0	0	0	+mo	0	$\Box(1,0)$	0	0
	ja	0	0	$h(en,+1)$	+ja	0	$\Box(1,0), \neg(2,+1)$	+1	+1
	ja	0	0	$h(en,+1)$	0	yes	$\Box(1,0), \neg(2,+1)$	+1	+1
	ja	0	0	$h(en,+1)$	0	0	$\Box(1,0)$	0	+1
	ja	0	0	0	+ja	0	$\Box(1,0), \neg(2,+1)$	+1	0
	ja	0	0	0	0	yes	$\Box(1,0), \neg(2,+1)$	+1	0
	ja	0	0	0	0	0	$\Box(1,0)$	0	0
q	ja	0	0	0	0	0	$\neg(1,0)$	0	0

도면8

source	jamo	moc	moi	context	cont - jamo	end	target	koi	eni
r	ja	0	0	0	0	yes	$\circ(1,0), \downarrow(2,+1)$	+1	0
	ja	0	0	0	+mo	0	$\equiv(1,0)$	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	$\equiv(1,0), -(2,+1)$	+1	+1
	ja	0	0	h(en,+1)	+mo	0	$\equiv(1,0)$	0	+1
	ja	0	0	h(en,+1)	0	yes	$\equiv(1,0), -(2,+1)$	+1	+1
	ja	0	0	r(en,+1)	-mo	yes	0	0	+1
	ja	0	0	r(en,+1)	-ja	yes	$\equiv(1,0), -(2,+1)$	+1	+1
	ja	0	0	r(en,+1)	+mo	0	$\equiv(1,0)$	0	+1
	ja	0	0	r(en,+1), h(en,+2)	+mo	0	$\equiv(1,0)$	0	+2
	ja	0	0	r(en,+1), h(en,+2)	0	0	$\equiv(1,0), -(2,+1)$	+1	+2
	ja	0	0	r(en,+1)	+ja	0	$\equiv(1,0), -(2,+1)$	+1	+1
	ja	0	0	0	-mo, +j a	0	0	0	+1
	ja	0	0	0	-ja, +j a	0	$\circ(1,0), \downarrow(2,+1)$	+1	+1
s	ja	0	0	s(en,+1), i(en,+2), o(en,+3), n(en,+4)	0	0	$\wedge(1,0), \downarrow(2,+1),$ $\sqsubset(3,+2)$	+2	+4
	ja	0	0	i(en,+1), o(en,+2), n(en,+3)	0	0	$\wedge(1,0), \downarrow(2,+1),$ $\sqsubset(3,+2)$	+2	+3
	ja	0	0	s(en,+1)	-mo, +m o	0	$\asymp(1,0)$	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	s(en,+1)	0	0	$\wedge(1,0)$	0	+1
	ja	0	0	0	0	0	$\wedge(1,0), -(2,+1)$	+1	0
	ja	0	0	0	+ja	0	$\wedge(1,0), -(2,+1)$	+1	0

도면9

source	jamo	moc	moi	context	cont - jamo	end	target	koi	eni
s	ja	0	0	0	+mo	0	$\wedge(1,0)$	0	0
t	ja	0	0	i(en,+1),o(en,+2), n(en,+3)	0	0	$\wedge(1,0), \exists(2,+1),$ $\sqsubset(3,+2)$	+2	+3
	ja	0	0	0	+mo	0	$\Xi(1,0)$	0	0
	ja	0	0	h(en,+1)	+ja	0	$\sqsubset(1,0), -(2,+1)$	+1	+1
	ja	0	0	h(en,+1)	0	yes	$\sqsubset(1,0), -(2,+1)$	+1	+1
	ja	0	0	h(en,+1)	0	0	$\sqsubset(1,0)$	0	+1
	ja	0	0	0	+ja	0	$\Xi(1,0), -(2,+1)$	+1	0
	ja	0	0	0	+ja	0	$\exists(1,0), -(2,+1)$	+1	0
v	ja	0	0	0	0	yes	$\exists(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	0	$\exists(1,0)$	0	0
	ja	0	0	0	+mo	0	$\neg(3,0)$	0	0
x	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
	ja	0	0	0	+ja	0	$\neg(1,0), -(2,+1)$	+1	0
z	ja	0	0	0	0	yes	$\neg(1,0), -(2,+1)$	+1	0
	ja	0	0	0	0	0	$\neg(1,0)$	0	0

도면10

