

한국인을 위한 중국어 성조 평가 시스템

김무중*, 김효숙*, 김신주*, 강효원*, 권철홍**

* (주)언어과학 언어공학연구소

** 대전대학교 정보통신공학과

Chinese Tone Evaluation System for Korean learners.

MuJung Kim*, Hyo-Sook Kim*, SunJu Kim*, HyoWon Kang and ChulHong Kwon**

* Center for Linguistic Engineering, Eoneo Inc.

** Department of Information & Communication Engineering, Daejeon University.

*{donaldos,hyosook,sunju,kanghyowon}@eoneo.co.kr

**chkwon@dju.ac.kr

Abstract

This study is about Chinese tone evaluation system for Korean learners using speech technology. Chinese pronunciation system consists of initials, finals and tones. Initials/finals are in segmental level and tones are in suprasegmental level. So different method could be used assessing Korean users' Chinese. Differ from segmental level recognition method, we chose pattern matching method in evaluating Chinese tones. Firstly we defined speakers' own speech range and produced standard tonal pattern according to speakers' own range. And then we compared input patterns of users with referring patterns.

I. 서론

본 연구는 음성기술을 응용하여 중국어의 발음체계 가운데 특히 성조 학습 시스템을 구성하는 방법에 관한 것이다. 중국과의 경제, 정치, 사회, 문화 등의 교류가 활발해지면서 중국어 학습 수요도 폭발적으로 증가하고 있다. 교육 방식도 문법 또는 작문 위주의 학습보다는 회화 중심의 학습으로 이루어지고 있다. 그러나 한국인들이 갖고 있는 한자 지식 능력에도 불구하고

하고 한국어와 중국어는 계통적, 구조적으로 전혀 다른 언어체계이므로 한국인이 중국어 발음을 습득하는 데는 많은 어려움이 있다[1].

특히 각 음절마다 다른 음높이로 실현되어 의미를 구별하게 해 주는 성조는 한국인이 학습하기에 매우 어려운 부분이다. 본고에서 표준 중국어로 설정한 普通话에는 5개의 성조가 있다(1-4성, 경성). 성조는 분절 음 층위에 해당하는 성모 및 운모와 마찬가지로 중국어 발음체계에서 매우 중요한 부분이다.

본 연구소는 음성기술을 이용한 언어교육(STILL: Speech Technology in Language Learning)을 주요 연구 주제로 삼고 있으며, 기 개발된 것으로는 한국인의 영어 발음 평가시스템이 있다. 외국어 학습 시스템을 구현함에 있어 특히 모국어의 간섭현상을 반영하는 대 중점을 두고 있다. 이를 이용하여 현재 중국어 발음 교정 시스템을 개발하는 과정에 있다.

II. 중국어 성조 평가 시스템

1. 구성

성조 평가 시스템은 그림 1, 2와 같이 크게 두 가지 하위 프로세스로 구성되어 있다. 그림 1에 도시된 첫 번째 프로세스는 화자별로 차이가 나는 음역을 상대화 하는 부분으로서 화자별 주파수 대역을 측정하여 이를 오도(五度)로 전환한다. 즉 화자의 음역에서 생성할 수 있는 표준 성조 패턴을 추출하는 부분이다[1]. 그림 2에 도시된 두 번째 프로세스에서는 학습자의 실제 성

조 패턴과 학습자의 음역을 기준으로 했을 때의 표준 성조 패턴을 비교한다. 즉 첫 번째 프로세스에서는 화자별 성조 참조 모델을 생성하며, 두 번째 프로세스에서는 학습자의 성조를 평가한다.

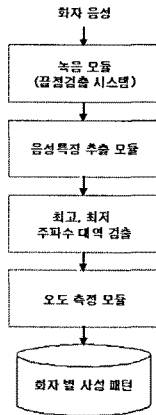


그림 1. 화자의 음역을 기준으로 한 참조 패턴 생성부

화자별 음역 결정 및 표준 패턴을 생성하기 위한 음성특징으로 주파수 도메인에서 분석한 F0와 시간 도메인에서 추출한 Pitch를 활용하였으며, 오도의 표현은 선형적으로 정의한 방법과 대수로 변환하여 정의한 방법을 모두 활용하였다. 그리고 화자별 표준 성조 패턴은 각 성조별로 간단한 수식을 통해서 생성하였다.

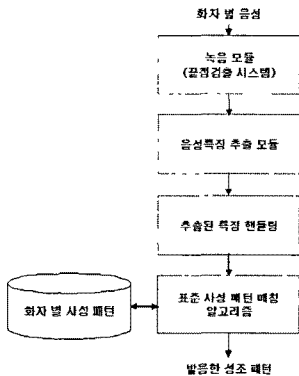


그림 2 참조 패턴과 입력 패턴의 비교를 통한 성조 평가부

학습자의 중국어 성조를 평가하기 위한 음성특징으로 중국어 성조 참조 패턴 생성과정에서 F0와 Pitch를 활용하였으며, 패턴 매칭 알고리즘으로는 DTW 기법과 선형비교법을 활용하였다. 이상의 두 가지 패턴 매칭

방법을 이용하여 참조 패턴과 입력패턴 간의 유사도 값을 출력하여 이를 토대로 입력패턴의 정확성을 평가하였다.

2. 음성 특징 추출 모듈

중국어 성조 인지 및 교육에 대한 연구에서 음성특징 파라미터로 음높이를 결정짓는 F0와 Pitch를 많이 활용하고 있다. F0와 Pitch는 동일하게 취급되지만, Pitch는 인지적인 측면에서의 양이고, F0는 물리적인 측면에서의 양이다.

본 시스템에서는 주파수 도메인에서 500Hz 이하에서 나타난 첫 번째 하모닉 성분을 tracking한 F0와 시간 도메인에서 auto-correlation 기법으로 음성의 주기성을 계산하여 추출한 Pitch를 음성특징 파라미터로 선정하여 활용하였다[2].

3. 오도 측정 및 정의 모듈

발화자 음성의 최고 또는 최저 값을 기반으로 오도를 정의하는 모듈이다. 화자의 대역은 화자의 '도,레,미,파,솔'에 각각 1도에서 5도까지를 대입하여 일차로 정의하였다. 이러한 방식에서는 다음과 같은 문제점이 있었다. 특히 5도 대역에서 화자발화가 불안정하였고, 1도 정도 낮게 발화되기도 하였다. 이는 로그 값을 이용하여 대역을 보완함으로써 해결하였다.

남성 발화자의 경우는 1도가 평균적으로 95Hz를 나타냈으며 5도는 평균적으로 155Hz를 나타냈다. 또한 여성 발화자의 경우에는 1도가 평균적으로 145Hz를 나타냈으며, 5도의 경우는 평균적으로 260Hz를 나타냈다.

4. 화자별 표준 성조 패턴 생성 모듈

중국어의 성조는 그림 3과 같은 형태로 구성되어 있다.

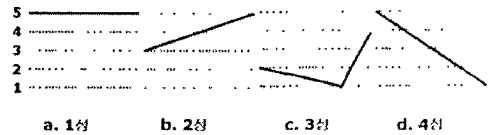


그림 3 중국어 사성 표준 패턴

1성은 가장 높은 5도에서 평탄하게 발화하는 형태이고, 2성은 3도에서 5도로 상승하는 형태이며, 3성은 2도에서 가장 낮은 1도로 하강 한 후 발화 구간의 75% 지점에서 4도로 상승하는 형태이다. 마지막 4성은 5도에서 1도로 하강하는 형태를 지닌다 각각의 성조 패턴은 간단한 수식으로 표현이 가능하다.

5. 추출된 음성특징 핸들링 기법

학습자의 중국어 성조 패턴과 학습자별 표준 참조 패턴을 매칭하기 전에, 유사도 측정 시에 발생할 수 있는 오류를 최소화하기 위하여 학습자의 음성을 다음과 같은 핸들링 기법을 이용하여 전처리해 주었다.

① Low Pass Filter 개념 활용

$$X = X_{t-1} + 2 * X_t + X_{t+1} / 4$$

② Median Filter

③ 대역 변화 점 추적 및 보정

6. 표준 성조 패턴과 학습자의 입력 패턴 매칭 기법

그림 1에서 도시한 중국어 성조 참조 패턴 생성부에서 생성한 참조 패턴과 입력한 화자의 성조 패턴의 유사도 계산을 위해 DTW 매칭 기법과 선형 매칭 기법을 이용하였다. DTW 매칭 기법은 참조패턴 생성 후 학습자의 성조패턴과 비교하는 기법이며, 선형 매칭 기법은 참조 패턴과 입력 패턴의 발화 데이터 길이의 차이를 보완해주는 기법이다.

III. 실험 및 결과 분석

본 장에서 실험은 6개월 이상 중국어를 학습한 경험이 있는 남녀 각각 10명의 200개 데이터에 대하여 분석한 결과이다. 1인당 발화 단어 수는 성조 별 각각 5개 단어씩 20개 단어를 발화하였다.

1. 음성특징(F0, Pitch)별 패턴 매칭 결과분석

일반적으로 F0와 Pitch는 동일한 의미로 많이 활용되나, 본 논문에서 Pitch는 시간 도메인에서 음성 파형의 주기성을 검출하여 이를 역으로 취한 주파수 값을 의미하며, F0는 FFT변환을 통해 500Hz 이하 대역에서 첫 번째 하모닉 성분을 tracking 한 음성특징을 의미한다[2][3].

위 실험 분석에서 패턴 매칭 기법은 DTW 알고리즘 그리고 5도는 대수 변환 형태로 정의한 대역을 활용하여 비교, 분석하였다.

1성, 2성, 4성은 각각 정확한 발음에 대하여 F0와 Pitch 각각에서 유사도 값이 차이가 나지 않는 반면, 3성에서는 F0를 음성특징으로 매칭하였을 때가 피치를 통해서 추출된 음성특징으로 매칭하였을 때보다 유사도가 높게 나오는 것을 알 수 있다. 이는 화자별 오도 측정 시, 그리고 학습자의 음성데이터에서 음성특징

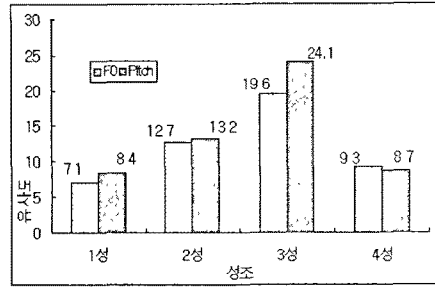


그림 4. 음성특징 별 패턴 매칭 유사도 추출 결과

추출 시 시간 도메인에서 추출한 Pitch 계산에서 Pitch 에러가 발생하였기 때문이다. 즉 주파수 도메인에서 추출한 F0가 시간 도메인에서 추출한 Pitch보다 안정적인 데이터 형태를 나타낸다고 할 수 있다.

2. 오도 대역 정의의 기법에 따른 패턴 매칭 결과분석

참조모델을 생성할 경우 오도대역 생성방법에 따라, 동일한 패턴이라도 유사도 평가 모듈에서 결과가 상이하다. 다음 그림은 대역 정의별 유사도 분석 결과이다. 실험 데이터는 음성특징별 패턴 유사도 분석 결과와 동일하며, 음성특징은 F0 그리고 매칭 기법은 DTW를 활용하였다.

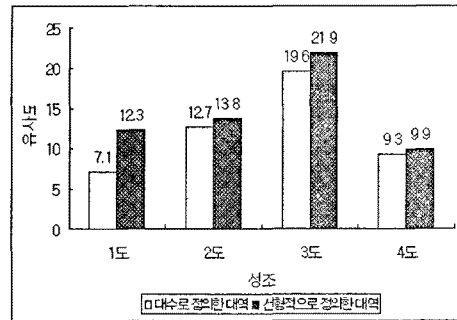


그림 5. 오도 정의의 대역별 패턴 매칭 결과분석

위의 결과에 따르면 1도와 2도가 각각 대역 정의의 별로 차이가 발생함을 알 수 있다. 이는 고주파 대역의 불안정한 발화 현상 때문이다. 한 예로 5도가 145Hz 정도가 되는 어떤 남성화자가 1성을 매우 불안정하게 발화할 수가 있다. 이런 경우에는 선형적인 대역 정의에서는 발화자의 오도 단계의 차이가 심하지만, 대수로 정의한 대역에서는 발화자와 참조데이터간의 데이터 값 차이가 선형적일 때보다 작게 나타난다.

3. 패턴매칭 기법별 매칭결과 분석

패턴 매칭 기법으로 DTW와 선형 매칭 기법을 적용

하여 분석하였다. 분석 방법으로는 각 성조별 추출되는 유사도 결과의 분산형태를 기준으로 하였다.

1,2,4 성에서 유사도 평가결과, 선형 매칭 기법의 유사도 결과의 분산이 작았으나, 3성에서는 DTW와 선형 매칭 기법이 1,2,4성에서의 분산 결과보다 크게 나타났다. 3성의 경우는 발화 구간 내에 음성특징 변화시점이 학습자의 패턴에서 매우 다양하게 나타났기 때문이다. 3성의 참조 패턴의 변화시점은 발화구간의 75% 지점으로 선정하였으나, 학습자 별로 50% - 90% 처럼 다양한 변화시점을 나타냈다.

IV. 결론

중국어 성조를 평가하기 위해서는 화자마다 차이는 나는 음역을 결정하는 것이 필요하다. 본고에서는 이를 위해 화자마다 음역을 정의하는 방식과 화자의 음역에서의 표준 성조 패턴을 추출해 내는 방법을 고안해 내었다.

중국어 성조를 표현할 수 있는 음성 파라미터로 시간 변이과정에서 안정적으로 추출할 수 있는 F0를 선정하였다. 특히 한국인이 중국어 사성 발화에서 4, 5도 대역에서 나타나는 1성의 오류 판정을 방지하기 위하여 대수변환 대역을 정의하는 것이 중요하다.

그리고 이러한 참조 패턴과 화자의 실제 발화를 비교하여 성조를 평가하였다. 특히, 3성 발성의 정확한 평가를 위하여 표준 중국어의 하강부분과 상승 부분의 변화시점을 통계적으로 분석하여 시스템에 적용할 필요가 있다[4].

추후 연구과제는 시스템의 평가와 중국어 음성학 전문가 집단의 평가 간의 상관관계를 측정하여 시스템의 안정성을 확보하고 성능을 개선하는 것이다.

감사의 글

본 연구는 정보통신부의 정보통신 산업경쟁력 강화 사업의 지원으로 수행되었습니다.
(과제번호 : A1200-0400-0013)

참고문헌

- [1] 심소희, “중국어 성조의 인지와 교육”, *말소리 제40호*, pp.67-77, 2000.
- [2] Stephen W.K.Fu, C.H.Lee, Oliville L. Clubb, “A Survey on Chinese Speech Recognition”, 1996.

[3] 신지영, *말소리의 이해*, 한국문화사, 2000.

[4] 송현선, “한국인 학습자의 2성 3성과 성조교육”, *중국 언어 연구 제 16권*, pp483-486, 2003.