

## 서울말과 방언사이의 억양차이 파라미터 추출

이강희<sup>o</sup>

<sup>o</sup>숭실대학교 글로벌미디어학부

e-mail:kanghee.lee@hcu.ac.kr<sup>o</sup>

## Extracting Speech Parameters for intonational Differences between the Seoul Dialect and the other Dialects of Korean

Kang-Hee Lee<sup>o</sup>

<sup>o</sup>Global School of Media, Soongsil University

### ● 요 약 ●

음성 인식 기술은 상당 기간 연구 개발이 계속적으로 되었는데 최근에 이르러 스마트 폰이 급속히 확산되면서 그 필요성, 즉, 고품질의 상용 서비스에 대한 욕구가 널리 확산되고 있다. 이런 환경에서도 사실 한국어는 상대적으로 소홀히 다뤄 질 수 있는 가능성이 매우 높다. 이러한 소외는 기술적인 문제로만 남겨지는 것이 아니라 언어라는 측면에서 문화와 직결된다. 이에 한국어 음성 인식 연구는 꼭 필요한 것이고 많은 부분 국가가 정책적으로 지원을 하는 것이 마땅하나 현 상황은 많이 미흡하나 아마도 곧 그 필요성이 대두 될 것이라 예상하며 그를 준비하는 연구로 특화된 분야, 즉, 표준어와 방언들 그리고 감정 표현 언어에 관한 연구를 한다.

**키워드:** 서울말(seoul dialect), 방언(dialect), 파라미터(parameter)

### I. Introduction

각종의 편리함과 여러 재미를 주고 있는 스마트폰 era에서 많은 응용 applications들이 있지만 그 중에서 인간의 일상생활에 가장 빨리 다가서고 있는 applications들 중 하나는 아마도 음성인식 시스템 이 아닌가 생각한다. 인간의 언어를 인식하여서 text로 바꿔주는 Speech-To-Text (STT) 혹은 conversely Text-To-Speech (TTS) 이거나 또 언어 그 자체의 뜻을 device로 하여금 수행시키는 음성인식 기술이 IT Industry의 Next Technology로 등장했다. 이것은 아마도 2011년 Apple의 아이폰 4s에서 SIRI 라는 음성인식 system이 소개되면서 관심을 받게 되었고 현재는 각 smartphone 제조사들이 여러종류의 음성인식 systems들 - 옵티머스 Q Voice, 갤럭시 S4의 S Voice, 그리고 베가 아이언의 Smart Voice 등 - 나오면서 그 경쟁이 가속되고 있다 [1]. 사실 Smartphone은 PC와는 달리 문자 입력이 불편하고 많은 경우 이동 중에 사용하기에 손을 사용하지 않는 다른 형태의 입력 기능이 필요하다. 이런 환경을 기반으로 음성 인식 기술이 이런 기능을 제공할 수 있다. 이에 한국어 음성 인식 연구는 꼭 필요한 것이고 많은 부분 국가가 정책적으로 지원을 하는 것이 마땅하나 현 상황은 많이 미흡하나 아마도 곧 그 필요성이 대두 될 것이라 예상하며 그를 준비하는 연구로 특화된 분야, 즉, 표준어와 방언들 그리고 감정 표현 언어에 관한 연구가 필요한 시점이다.

### II. Preliminaries

음성 인식 기술은 그 특성상 언어별로 기술 개발이 필요하다. 물론, 그 핵심 알고리즘은 큰 차이가 없으나, 언어별로 최적화 과정이 필요하며 더불어 Acoustic Model, Language Model, Word Model, 등 특정언어에 맞는 별도의 기술개발이 필요하다 [2][3]. 따라서 이 기술의 투자와 연구 개발은 주로 사용자가 많은 특정 언어들 (영어, 중국어, 등) 에 집중될 것이기에 한국어와 같이 사용자가 상대적으로 소수인 경우에는 소외될 우려가 있다. 실제로 마이크로 소프트웨어의 software products에 장착된 음성 인식 기능 - 특별히, 장애인을 위한 - 이 영어, 중국어, 일본어만 지원하고 있으며, Nuance Communications의 음성 인식 softwares 들도 주로 영어, 중국어, 독일어, 스페인어, 일본어 등만 지원하고 있다. 이러한 소외는 기술적인 문제로만 남겨지는 것이 아니라 언어라는 측면에서 문화와 직결된다. 이에 한국어 음성 인식 연구는 꼭 필요한 것이고 많은 부분 국가가 정책적으로 지원을 하는 것이 마땅하나 현 상황은 많이 미흡하나 아마도 곧 그 필요성이 대두 될 것이라 예상하며 그를 준비하는 연구로 특화된 분야, 즉, 표준어와 방언들 그리고 감정 표현 언어에 관한 연구를 한다.

### III. Proposed Scheme

음성 인식 기술을 간단히 보면 입력 음성을 computer가 분석하고 그에 관련된 특성 parameters들을 추출한 다음, 미리 수집된 데이터베이스들과 비교하여 가장 유사한 것을 그의 결과로 나타내는 일련의 과정이다. 그 typical한 블록 다이어그램을 다음의 그림 2에서 보여준다.

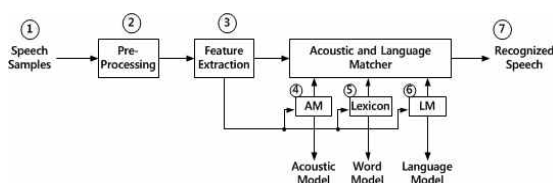


Fig. 1. A Typical Block Diagram of a Speech Recognition System

이외에도 accents, intonation patterns, pauses, duration, 등이 있다. Fig. 1에서 3의 Feature Extraction에서 본 연구에 필요한 여러 parameters들을 추출한다. 앞에 언급한 여러 measurements들에서 본 연구에서는 중요한 몇가지의 parameters들을 - pitch range, average pitch, speech rate, F0 contour, Formant frequencies - 등을 조사해보고, 많은 연구들에서 잘 하지 않았던 higher order spectrum을 알아보고자 한다.

### IV. Conclusions

본 논문의 서울말과 방언사이의 억양차이의 파라미터를 추출한다면 자연어 인식 분야에서 한국어의 독보적인 인식기술 이니셔티브를 확보할 수 있을 것이다.

### Acknowledgement

이 논문은 2014년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 연구되었음(NRF-2014R1A1A1A05008028)

### References

- [1] Comparison of Vocie Recognition according to the Smartphone Manufacturers, <http://www.playwares.com/xe/32740832>
- [2] Janet E. Cahn, "Generating Expression in Synthesized Speech", Technical Report, M.I.T. Media Lab., 1990.
- [3] Chin-Hui Lee and Sabato Marco Siniscalchi, "An information-Extraction Approach to Speech Processing: Analysis, Detection, Verification, and Recognition", Proceedings of the IEEE, Vol. 101, No. 5, pp. 1089-1115, May 2013.