

한국어 서술어의 문장말 위치에서의 의양패턴에 대한 자동인식

Automatic Recognition of Sentence-final Developmental Patterns for Korean Predicates

이 기 영¹, 송 민 식², 최 창 식³
 1 연세대학교 대학원 언어학 석사과정
 2 연세대학교 대학원 컴퓨터언어학 석사과정
 3 연세대학교 대학원 컴퓨터언어학 석사과정

요 약

최근, 문장단위의 음성을 인식할 수 있는 시스템을 개발하는 단계에 접어들면서 자발적인 발성음성의 인식 또는 음성언어 이해의 차원을 위한 시스템의 개발을 위해 문음특징을 이용하는 연구가 요구되고 있으나, 지금까지 개발되어 온 음성인식시스템은 주로 독립단어의 인식수준에 머물고 있기 때문에 문음을 이용하고자 하는 연구가 상대적으로 미흡한 수준에 있다.

본 연구에서는 한국어의 중의성(ambiguity) 문장에서 서술어 부분을 세그먼트하고 이 부분의 의양패턴을 자동인식하여 중의성 문장이 서술형, 의문형, 명령형, 권유형인지를 파악함으로써 인식시스템에서 의양패턴을 이용할 수 있는 가능성을 제시하였으며, 서술형 문장음성의 서술어 부분의 의양변환에 의해 의문형, 명령형, 권유형 문장으로 변형시키므로써 서술어 부분의 의양패턴에 따라 문장의 형태가 구분될 수 있음을 확인하였다.

차원에서 더욱 정확하게 음성의 이해나 자동인식을 수행할 수 있게 되며, 음성처리분야의 합성분야에서는 이미 주요한 제어요소로서 자리잡고 있기는 하나, 보다 자연스러운 음성의 처리과정을 위해서는 필수불가결한 특징패턴으로서의 역할을 하게 될 것이다.

문자로 쓰여진 문장에서는 많은 부분에서 총사구조 또는 의미구조상 여러가지 의미, 즉 중의성(ambiguity)이 존재한다[7]. "학교에 가요"라는 한 문장은 다음과 같은 억양에 의해 네가지 의미로 사용될 수 있다[8].

- a. : 서술형
- b. : 의문형
- c. : 명령형
- d. : 권유형

이억양패턴을 자동으로 인식할 수 있는 시스템을 개발한다면 네가지 의미중에서 올바른 것을 파악하여 자연언어 처리분야에 제공함으로써 시스템의 효율을 높일 수 있다.

본 연구에서는 한국어의 문장말 서술어에 대한 의양패턴을 작성하고 이를 자동인식하므로써 중의성 문장의 의미가 서술형인지 의문형인지 명령형인지 권유형인지를 구분할 수 있는 방법을 제안하며 작성된 의양패턴을 이용하여 중의성 서술어의 억양을 변환하므로써 문장의 의미를 전달할 수 있는 적절한 패턴인지를 확인하였다.

2. 서술어의 의양패턴

발전된 문장의 의미를 가장 확실하게 나타내는 특징은 서술어의 억양이라고 할 수 있다. 문장을 서술형, 의문형, 명령형 또는 권유형으로 나누면 억양으로 나타나는 F0의 궤적이 각각 서로 다르게 나타난다. 따라서 서술어의 의양패턴을 작성하기 위해 먼저 문장의 서술어 부분을 세그먼트하였으며 이부분의 억양특징을 추출하였다. 그림 1에 서술어의 의양패턴을 작성하기 위한 불력도를 보이고 있다.

1. 서론

한국어를 포함한 대부분의 언어에서 음절보다 더 큰 단위의 구절이나 문장 전체에 걸쳐 총사구조적 대립을 이루는 요소인 억양이 존재하며, 문장 내에서는 화자의 의도에 따라 필요한 강조부분이나 서술어의 높낮이 등이 억양에 의해 표현된다. 억양은 같은 내용의 음성을 발성한다 하더라도 발성된 음성의 구분이 서술형인지, 의문형인지, 명령형인지 아니면, 권유형인지에 대한 정보를 제공할 뿐만 아니라, 음성을 발성한 화자의 개인성, 정서성 및 심리적 인상등을 제공하므로 억양은 음성에 의해 전해지는 매우 중요한 정보이다. 어휘나 음성의 각 음절들을 정확히 발음하기 보다는 발성할 문장 전체의 구문형식에 따른 억양의 의미의 면저 배우며, 음절들의 정확한 청취보다는 억양의 변화에 대한 청취에 의해 발성문장의 기본적인 뜻을 파악하게 된다. 그러나, 지금까지 대부분의 음성자동인식에서는 입력된 음성의 단어에 대한 음향적 특징벡터들의 상관관계나 외곽의 정도만을 이용함으로써 동일한 단어로 구성된 문장이면 억양에 의해 전달되는 발성화자의 의도와 관계없이 동일한 음성으로 인식해 버리는 등의 문제를 안고 있다[1, 2].

언어학계에서는 억양에 대해 음성학적으로 많은 연구가 있어 왔지만 최근의 음성언어 인식에 관한 연구에 그리 큰 영향을 미치지 못하고 있다. 아직까지 음성이거나 인식분야에 억양패턴을 이용하지 못하고 있는 주요한 인인은 발성된 음성으로부터 억양을 암시하며 음성이거나 인식분야에 적절히 사용할 특징을 자동적으로 추출할 수 있는 정확한 방법이 제시되지 못하고 있기 때문이다[3, 4].

음성으로부터 음성처리분야에 적절한 억양패턴이 생성되면 전체의 음성언어와 자연언어처리[5, 6]를 보다 바람직한 방향으로 유도할 수 있을 뿐만 아니라, 음성 자동인식에서 동일한 단어로 구성된 음성이라 하더라도 억양패턴을 인식함에 의해 단어수준에서의 음향적 특징벡터의 상관관계나 외곽을 특징하여 구문적이나 의미적으로 분류를 비롯 내외적인 문제를 피하고 보다 높은

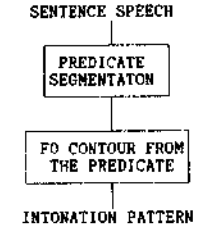


그림 1 억양패턴의 작성

그림 1에서 서술어 부분의 세그먼트를 구하기 위하여 에너지 특징을 이용하였으며, F0를 추출하기 위하여 center clipping and autocorrelation [9]을 이용하였다.

3. 억양인식과 변환

3.1 특징추출

문장음성의 서술어 부분을 세그먼트하고 억양특징을 추출하기 위하여 에너지와 F0의 궤적을 이용하였다. 음성신호의 샘플링 주파수는 10kHz로 하였으며, 프레임 길이는 256샘플, 간격은 128샘플로 하였다. 한국어 문장의 서술어 부분이 문장의 마지막 부분에 위치하였으며, 억양을 가진 유성음을 포함하고 있으므로 서술어 부분을 세그먼트하기 위하여 문장음성의 끝점에서 거꾸로 올라오면서 에너지가 떨어졌다가 올라오는 부분까지 문장음성의 서술어 부분으로 정하였고 추출한 서술어 부분의 억양특징은 F0의 궤적으로 하였다. 그림 2는 중의성 문장음성의 4가지 형태에 따른 음성파형과 각각 세그먼트되는 서술어 부분의 에너지와 F0의 궤적을 보이고 있다.

3.2 자동인식

문장음성의 서술어부분은 그 발성길이 각각 다르므로 억양패턴의 길이도 달리 표현되며, 권유형 서술어의 발성길이는 보통 명령형 서술어의 발성길이의 1.5배 이상이므로 비선형시간축정규화기법을 이용할 필요가 있다. 따라서, 억양특징으로부터 구성된 억양패턴을 자동인식하기 위하여 DTW를 이용하였다. 또한 억양패턴 사이의 거리는 다음 식으로 정의 하여 이용하였다.

$$d(F_i, F_j) = |F_i - F_j| \quad (1)$$

$$F_i = F_{0i} / \bar{F}_0 \quad (2)$$

여기서 F_i 는 추출된 F_{0i} 를 평균치로 나눈 것이다.

3.3 억양변환[10]

중의성 문장의 서술어의 억양을 변환시키기 위하여 3.1절에서 세그먼트된 서술어 부분의 여기신호를 이용하였다. 미량법에서는 서술어가 발생되는 지속시간이 변환시키고자 하는 여기신호와 길도 변환한다.

4. 실험

4.1 음성데이터베이스

본 실험에서는 중의성 서술어에서 4가지 형태의 서로 다른 억양패턴을 선택하기 위하여 표 1과 같은 중의성 문장을 선택하였다. 서술어의 유성음 부분에서 F0의 궤적을 추출하였지만 음성에 따라 F0의 궤적의 연속성이 없거나 부분적으로 연결하여 억양패턴을 작성하였다.

표 1 중의성 문장

(반말)
(학교에 가, 여기에 서)
(반문장)
(학교에 가요, 여기에 서요)
(중칭)
(학교에 가세요, 여기에 서세요)

본 실험의 표 1과 같이 중의성 문장의 종류는 6문장이며 서술어의 억양에 따라 4가지 형태의 문장이 되므로 24 문장을 실험대상으로 하였으며 2회자가 2번씩 발생하였으므로 실험에 쓰인 문장은 총 96 문장이다. 억양패턴의 표준패턴은 2회자가 첫번에 발생한 6문장에서 각 서술어 부분의 억양패턴인 12개의 다수 패턴으로 구성되며 나머지 억양패턴을 시험패턴으로 하였다.

4.2 인식결과 및 고찰

4.1절의 시험패턴과 표준패턴에 의해 식(1)의 거리를 이용하여 자동인식을 수행하였다. 표 2는 중의성 문장서술어의 특정화자별 4가지 억양패턴에 따른 인식결과를 나타내었으며, 표 3에는 불특정화자별 4가지 억양패턴에 따른 인식결과를 나타내었고, 표 4에는 지속시간을 고려한 경우의 불특정화자별 인식결과를 나타내었다.

표 2 억양패턴의 인식결과

(a) 특정화자

	서술형	의문형	명령형	권유형	평균
인식율	91.6	100	25.0	50.0	66.7

(b) 불특정화자

	서술형	의문형	명령형	권유형	평균
인식율	91.6	100	25.0	50.0	62.5

표 2에서 인식율이 제일 낮은 문장의 경우 명령형과 권유형이었다. 명령형과 권유형의 억양패턴은 전체적으로 높은 유사한 모양을 갖지만 엄격히 다른 특징은 각각의 지속시간이므로 표 3에는 명령형과 권유형의 오인식이 발생할 때 두차례로 지속시간을 고려한 인식결과를 나타내었다.

표 3 지속시간을 고려한 억양패턴의 인식결과

(a) 특정화자

	서술형	의문형	명령형	권유형	평균
인식율	91.6	100	91.6	75.0	89.5

(b) 불특정화자

	서술형	의문형	명령형	권유형	평균
인식율	83.3	100	75.0	66.7	81.2

지속시간을 고려함으로써 인식율을 약 20% 향상시킬 수 있었다.

4.3 억양변환결과

3.3절의 억양변환에 의해 서술형 문장의 서술어 부분을 억양변환한 결과 의문형, 명령형 및 권유형의 문장으로 변환시킬 수 있었다. 그림 3은 억양변환에 의한 음성파형 및 그 특징들이다.

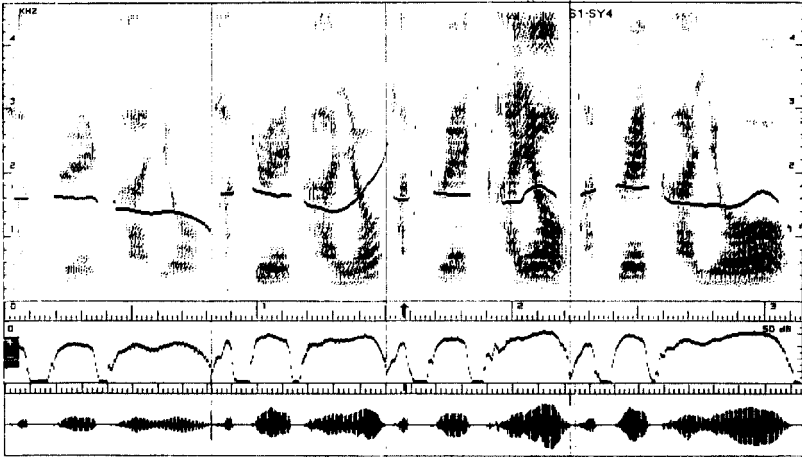


그림 2. "애", "이", "에", "우"의 4가지 모음의
 (서술형, 이음형, 이음형, 이음형)

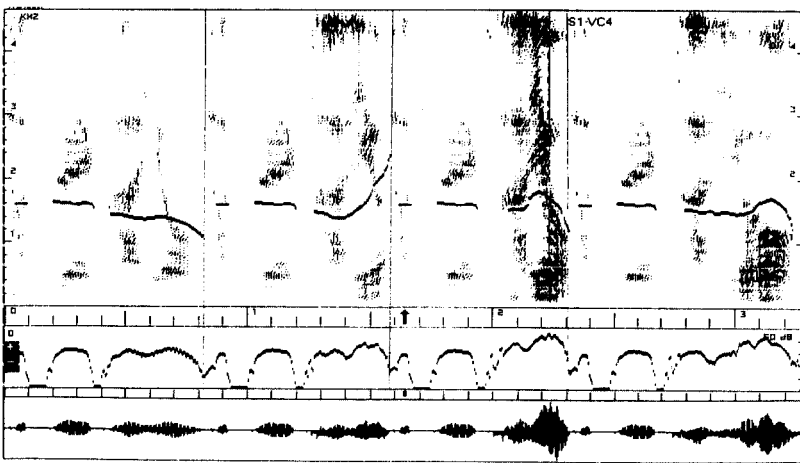


그림 3. "애", "이", "에", "우"의 4가지 모음의
 (서술형, 이음형, 이음형, 이음형)

5. 결론

본 연구에서는 문장음성의 서술어에 대한 억양패턴을 작성하기 위하여 에너지와 F0의 궤적을 이용하였으며 추출된 억양패턴을 자동인식하기 위하여 비선형정규화 기법인 DTW를 이용하였다. 중의성 문장들을 대상으로 실험한 결과 다음과 같은 결론을 얻을 수 있었다.

- (1) 에너지 궤적을 이용하여 서술어 부분의 세그먼트가 가능하였다.
- (2) 서술어에 대한 억양패턴을 자동인식하여 중의성 문장음성을 4가지 형태(서술, 의문, 명령, 권유)를 60% 정도 인식할 수 있었으며, 지속시간을 고려할 경우 80% 정도의 인식이 가능 하였다.
- (3) 서술어의 여기신호를 이용한 억양변환에 의해 중의성 문장을 4가지 형태(서술, 의문, 명령, 권유)로 변환할 수 있었다.

* References *

- [1] S. Furui, M.M.Sondhi, Advances in Speech Signal Processing, Marcel Dekker, Inc., 1992
- [2] W.A. Lea, Trends in Speech Recognition, Prentice-Hall Inc., 1980
- [3] J.W. Butzberger, Jr., et al., "Isolated Word Intonation Recognition Using Hidden Markov Models," in Proc. ICASSP 90, pp. 773-776, 1990
- [4] C. Wightman, M. Ostendorf, "Automatic Recognition of Intonation Features," in Proc. ICASSP 92, pp. I-221-4, 1992
- [5] 송민석, 최현락, 이찬도, 정국, "음성언어인식의 이론적모형," HCI, 1994
- [6] 고도홍, "음성합성시 PROSODY의 처리: 한국어의 INTONATION을 중심으로," 음성통신 및 신호처리 WORKSHOP, 한국음향학회, 1990
- [7] 정국, 윤율단위 음운론 및 음운통계에 관한 연구, 한국전자통신연구소 위약과제연구 최종보고서, 1993
- [8] 이기문, 외2인, 국어음운론, 학연사
- [9] L. R. Rabiner, R. W. Schafer, Digital Processing of Speech Signals, Prentice-Hall, Inc., Englewood Cliffs, N. J., 07632
- [10] 이기영, 최창식, "타회자의 여기신호를 이용한 억양변환," 한국음향학회 논문심사중, 1995