음성합성의 운율처리를 위한 악센트 연구

구 희 산
(중앙대학교)

이 논문은 음성합성에 있어서 운율처리와 합성 과정에 관한 소개와 함께 한국어 악센트 분석을 통하여 운율모델링에 기초자료를 제공하는 데 그 목적이 있다. 현 단계의 합성음의 품질은 분절음의 명료성에서 상당한 진전을 보이고 있으나 운율의 처리가 아직 미숙하여 전반적인 합성음의 자연성이 미달된다. 이에 대한 근본적 이유는 운율 특성에 대한 이해가 부족하기 때문이라고 본다. 따라서 본 논문에서는 표준한국어 악센트를 분석하고, 음성합성시 운율정보 추출의 문제를 제시하기로 한다.

1. 서 론

최근 인간과 컴퓨터 간의 상호 정보 전달의 필요성이 확대됨에 따라 인간의 음성과 같은 자연스런 합성음 개발에 관한 요구가 증대되고 있다. 음성합성이란 기계로 하여금 인간이 정보수단으로 사용하는 말소리와 유사한 소리를 만들게 하는 것이다. 현재 미국이나 일본에서는 30-40년간에 걸쳐 축적된 음향음성학 지식을 바탕으로 상당한 수준의 합성기술이 발전하고 있다. 한편 국내 여러 기술협력단체나 대학의 연구소에서도 이 분야의 기술 개발에 박차를 가하고 있는 실정이다. 그렇지만 국내에서 발표된 합성음의 품질은 외국것에 비하면 자연성에서 크게 못 미치고 있다. 이러한 원인 중의 하나가 바로 한국어에 대한 음향음성학 연구의 부족이라 할 수 있다. 특히 운율에 관한 기초자료가 부족하기 때문에 합성음에서 악센트나 억양과 같은 운율의 처리가 미흡하다. 그러므로 본 논문의 전개에 있어서 음성합성의 절차 중에서 운율정보에 관련되는 단계를 중심으로 피력하고, 필요에 따라 예를 들어 설명하기로 한다.

2. 음성합성

음성합성은 정서법으로 쓰여진 문장을 음성으로 변환하는 기술로 흔히 문장-음성변환 시스템이란 용어로 사용된다. 이 시스템은 도표 1처럼 크게 언어처리 기능과 음성합성 기능으로 구분되는데 자세한 것은 구성도를 보고 이해하기로 한다.

2.1. 언어처리 기능

언어처리 기능은 전처리, 구문분석, 운율정보 추출, 글자-음운 변환, 단위음성 생성으로 구성된다. 그리고 음성합성 기능은 단위음성 합성과 운율합성으로 구성된다.

(1) 전처리

외래어, 숫자, 약어 및 단위용어와 같은 특수한 내용을 일반 텍스트로 변환시키는 일이 전처리 과정에서 수행된다. 예를 들면 약어나 기호 사전을 참고로 'O.K'를 '오케이'로, 'cm'은 '센티미터'로 읽을 수 있도록 한글로 변환시켜 주는 과정이다.

Copyright (C) 2003 NuriMedia Co., Ltd.

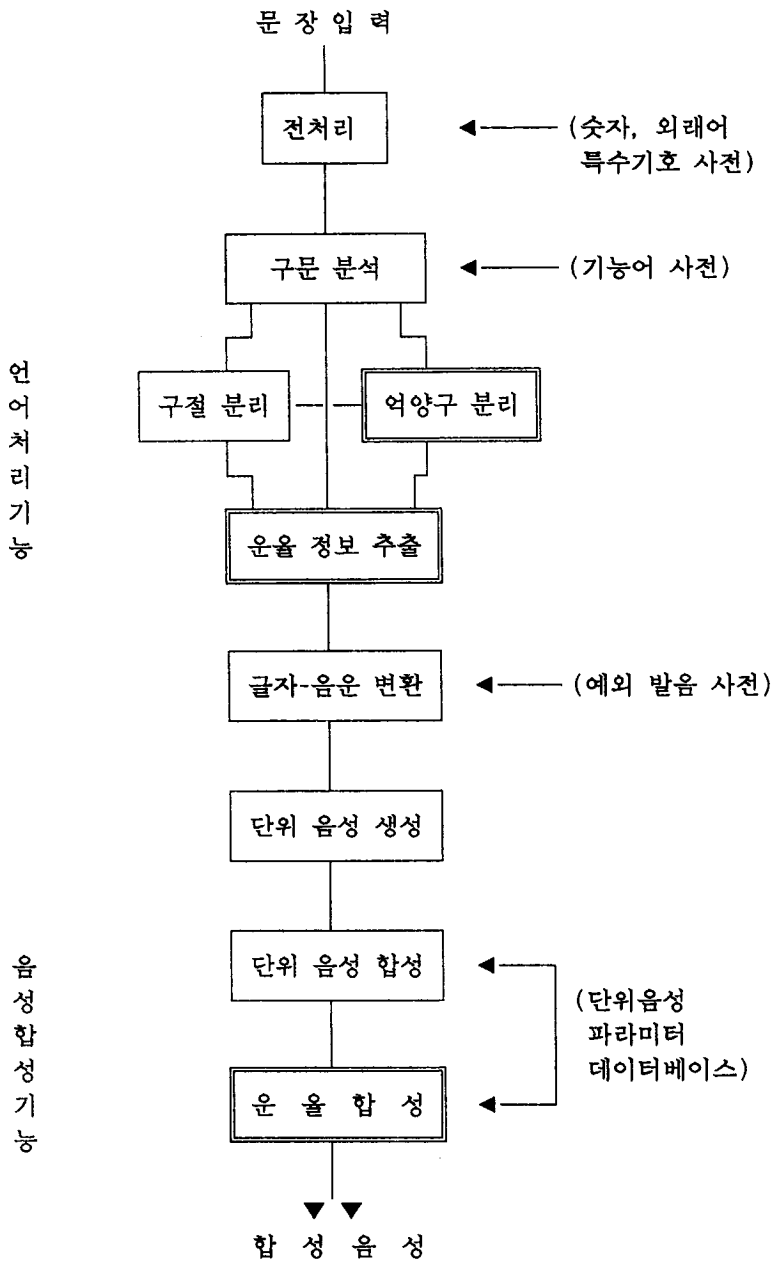


도표 1. 음성합성 시스템의 구성도

Copyright (C) 2003 NuriMedia Co., Ltd.

(2) 구문분석

이 과정은 입력된 문장의 형태론과 통사론적 분석으로 구분된다. 형태론적 분석은 어휘와 형태소 사전의 조합규칙을 이용하여 문장의 구절을 분석하는 것이고, 통사론적 분석은 통사규칙과 형태론적 분석의 결과를 기반으로 구절간의 문법적 결합관계를 분석하는 것이다. 구문분석 과정에서는 기능어 사전을 이용하여 주로 각종의 조사와 어미 등을 처리한다. 그리고 구절분리는 억양구분리와 밀접하게 연관되어 있어서 상호 필요한 정보를 사용하여야 한다. 더우기 억양구분리는 직접적으로 휴지나 악센트 위치 같은 운율정보를 제공해 준다.

(3) 운율정보 추출

운율은 발화단위에 내포된 의미정보에 따라 결정되는데 구와 절의 위치, 기능, 상호 결합관계 등의 복합적인 요인에 영향을 받게 된다. 한국어의 운율은 강세구조가 있는 영어처럼 복잡하지 않아서 주로 조사나 어미 등과 같은 기능어의 문법정보만으로도 구절단위의 운율정보를 얻을 수 있다. 구절간의 문법적 결합관계에 따라 억양구가 정해지므로 운율정보는 구문분석 단계와 밀접한 관계를 갖는다. 이 단계에서 추출해야할 자료는 구절과 억양구의 분절, 구절의 억양패턴, 악센트의 위치와 성분 그리고 구절의 경계인 휴지의 위치와 길이 등이다. 여기서 얻어진 운율정보는 특수한 기호를 이용하여 표기하고 이 기호들은 합성처리 단계에서 이용되어 운율합성기에서 실제 운율을 만들게 된다.

(4) 글자-음운 변환

입력된 문장을 분석한 다음 음성기호를 사용하여 각 어절의 글자를 소리나는대로 바꾸어 주어야 한다. 이 단계가 글자-음운 변환 단계인데 여기에서는 모든 음운변동 규칙이 적용된다. 음운변동 규칙에는 구개음화, 연음규칙, 경음화, 자음동화, 모음의 장음화 등이 포함된다. 특히 주의하여야 할 것은 규칙의 적용 순서를 따라야 하는 일이다.

(5) 단위음성 생성

단위음성 생성은 합성단계에서 접속할 음성을 미리 단위별로 저장하는 단계라 할 수 있는데 단위음성은 대개 음소, 반음절, 음절 중에서 선택된다. 각 단위는 장단점이 있어서 음소를 사용하면 자료의 양은 적으나 인접 음소와의 전이의 실현이 어렵고, 반음절은 음소단위보다는 자료가 많으나 스펙트럼상의 불연속을 최소화하여 음 사이의 전이를 유지할 수 있으므로, 현재 가장 많이 사용되고 있다. 그리고 음절을 사용하면 자료의 양이 많으며 변이음 처리에 어려움이 따른다(이양희, 1992).

2.2. 음성합성 처리 기능

언어처리 과정에서 저장된 운율정보 추출과 단위음성 생성의 자료를 합성처리 과정에서 넘겨받아 실제의 합성음을 만들게 된다. 합성법은 포먼트(Formant)합성방식과 분석합성(Vocoder)방식으로 크게 두가지로 분류된다. 분석합성방식으로는 LSP(Line Spectrum Pair), 켑스트럼(Cepstrum), 파코(Parcor), PSOLA(Pitch Syncronous Overlap and Add)등이 있다. 이 과정에서 처리될 일은 단위음성 생성 과정에서 구축된 데이타베이스를 이용하여 단위음성을 합성하고 자연성을 높이기 위해 운율을 합성하는 일이다.

Copyright (C) 2003 NuriMedia Co., Ltd.

(1) 단위음성 합성

실제 음성을 만들기 위해 데이타베이스에 구축된 단위음성이 이 단계에서 결합된다. 스펙트럼의 결합은 자동으로 이루어지도록 하여야 하며 불연속성이 최소화될 수 있게 하여야 한다. 단위음성 결합기술이 합성음질에 가장 큰 영향을 미치므로 몇가지 고려하여야할 사항은, 첫째, 적절한 분석용 발화단위 선택방법, 둘째, 정확한 반음절 경계 추출 및 분할방법, 세째, 두 반음절 단위의 연결방법, 그리고 네째, 음절길이 조절방법 등이다(안성권/성광모, 1992).

(2) 운율합성

운율은 운율조절기가 운율규칙에 따라 추출한 운율 파라미터로 조절된다. 음성합성에서 자연성을 높이는 세가지 성분은 기본주파수(Fo: Fundamental frequency), 강도(Intensity) 및 지속시간(Duration)이다. 기본주파수는 음의 높낮이를, 강도는 음의 강약을, 그리고 지속시간은 음의 길이를 각각 조절하는 기능을 한다. 운율처리는 언어의 특성에 따라 다르므로 해당 언어의 운율 성분 분석이 전제가 되어야 한다. 영어나 일본어의 경우 기본주파수 모델이 거의 확립되었으나, 한국어의 경우 이 분야의 연구가 미흡한 상태다. 한국어 악센트의 성분이나 위치가 구절의 분절에 필수적인 역할을 하며, 또한 구절의 분절은 결과적으로 억양구 분절과 억양패턴에 밀접히 연관되어 있으므로 기본주파수 뿐만아니라 강도나 지속시간도 악센트에 대한 분석을 통하여 모델화되어야 한다.

3. 악센트

일반적으로 악센트(Accent)라는 용어는 강세(Stress)와 혼용해서 사용되기도 하지만 주로 강세는 영어에서처럼 단어 자체 내에서 일정한 음절이 인접 음절보다 특정한 음향성분에서 우세하게 나타남을 표기하기 위해 사용되고, 악센트는 인용된 단어 자체 내에서 보다도 실제 발화에서 일정한 음절에 특정한 음향 성분이 다른 음절보다 더 뚜렷하게 표출될 때 사용된다. 이 논문에서 악센트란 바로 이러한 의미로 사용하기로 한다.

악센트의 음향성분은 고저(Pitch), 강도 및 지속시간으로 분류하여 측정한다. 그러나 악센트는 어느 한가지 성분만으로 구성되어 있는 것이 아니고 세가지 모두가 혼합되어 구성된다. 최근까지 악센트의 음운론적 연구에서는 세가지 성분 중에서 가장 뚜렷한 것이 무엇인가에 관심을 두고 단순성분론을 펴왔으나, 요즘 음성합성이나 인식과정에서는 더욱 세밀한 성분 분석이 요구됨에 따라 복합성분론에 관한 연구가 활발히 이루어지고 있다.

3.1. 단순성분 분석

(1) 고저와 기본주파수

고저음은 인간에게 지각되는 음의 높낮이를 뜻하고, 이것을 물리적 기구로 측정할 때는 기본주파수를 이용한다. 그런데 기본주파수가 높아지거나 낮아짐에 따라 어느 한계까지는 고저도 그만큼 변하지만 어느 한계를 지나면 상관관계가 약해진다. 왜냐하면 고저란 음조에 대한 심리적인 속성의 하나이고, 또한 인간의 생리구조가 무한한 기본주파수 영역을 다 들을 수 있게 만들어지지 않았기 때문이다. 그러므로 고저음은 음의 자극에 대한 직접적인 주관적 반응으로서 청자에 의해 판단되고, 반면에 기본주파

Copyright (C) 2003 NuriMedia Co., Ltd.

수는 물리적 도구에 의해 객관적으로 측정된다. 기본주파수가 연속적으로 이어질 때 곡선이 생기는데 이것을 기본주파수곡선이라 하며(도표 2-A 참조), 한편 억양연구 측면에서는 이것을 억양곡선이라 칭하여 결국 같은 것을 의미한다. 기본주파수의 측정 단위로는 Hz(Hertz)가 사용된다.

(2) 강도와 진폭(Amplitude)

운율을 분석할 때 음의 시끄러운 정도를 측정하는 것이 강도인데, 여기서 강도는 성대안의 성문이 열리는 정도와 성문을 통과하는 공기의 양과 속도에 따라 결정된다. 음성이나 기계소리같은 것을 객관적으로 측정할 때 쓰이는 물리적 용어는 진폭이다. 그리고 강도와 같이 쓰이는 또 다른 심리적 용어는 세기(Loudness)이다. 이때 강도나 진폭을 재는 척도로 dB(decibel)을 사용한다. 강도를 측정하는 방법은 다음과 같은 세 가지가 있다.

첫째, 정상강도(Peak intensity)의 측정: 악센트 받는 음절에 해당하는 진폭 중에서 진폭의 값이 가장 높이 나타나는 지점을 재는 것이다.

둘째, 전체강도(Total intensity)의 측정: 악센트 받는 음절에 해당하는 진폭의 총계로 전체의 에너지를 재는 것이다.

세째, 평균강도(Average intensity)의 측정: 전체강도를 악센트 받은 음절의 길이로 나누는 방법이다.

이상과 같은 측정 방법중 인접 음절간의 상대적으로 시끄러운 정도를 잴 때에는 일반적으로 두번째의 전체강도 측정법을 사용한다(도표 2-B 참조).

(3) 지속시간과 시간(Time)

지속시간이란 음성분석에서 쓰이는 용어로 발화 속에서 자음, 모음 또는 음절이 지속되는 길이를 말한다(도표 2-C 참조). 합성음에서 각 음소나 음절의 길이에 따라 다양한 음질이 결정되므로 지속시간의 처리가 대단히 중요하다. 특히 자연성과 밀접히 관련되는 음성의 속도, 즉 느리고 빠른 음성을 만들 때 각 음성 단위의 지속시간에 대한 정확한 정보가 요구된다. 지속시간은 발화 단위상의 음절의 위치에 따라 달라지므로 위치별로 길이에 대한 자료와 속도조절규칙이 데이타베이스에 저장되어야 한다.

현재 낭독체로 읽은 보통 말소리의 음소별 길이에 대한 연구는 되어 있으나 말소리 속도나 문체에 따른 연구는 미진한 상태다. 지속시간은 스펙트로그램에 나타나는 음소와 음절의 경계를 기준으로 측정할 때 파열음이나 마찰음의 경계는 에너지의 분포로 쉽게 확인이 되나 설측음, 비음 및 반모음 같은 공명음들의 경계는 스펙트로그램상에 명확하게 나타나지 않기 때문에 분절하는데 어려움이 따른다. 지속시간의 측정단위는 ms(Milli-second)를 사용하며, 이것은 초(Second)를 길이 단위로 나눈 것이다.

3.2. 복합성분 분석

(1) 고저음과 강도

Stevens(1935)와 Snow(1936)는 높은 소리의 고저는 강도와 함께 높아지고, 한편 낮은 소리의 고저는 강도와 함께 낮아진다고 보고하였다. Terhardt(1975)도 이것과 유사한 결과를 보고하였지만 인간의 말소리같은 복합음조(Complex tone)는 저음에서 약간의 부정적인 영향을 보이기도 한다고 주장하였다. 그렇다면 고저의 범위에 따라 고저음과 강도는 항상 상관관계가 있다고 볼 수는 없을 것이다.

Copyright (C) 2003 NuriMedia Co., Ltd.

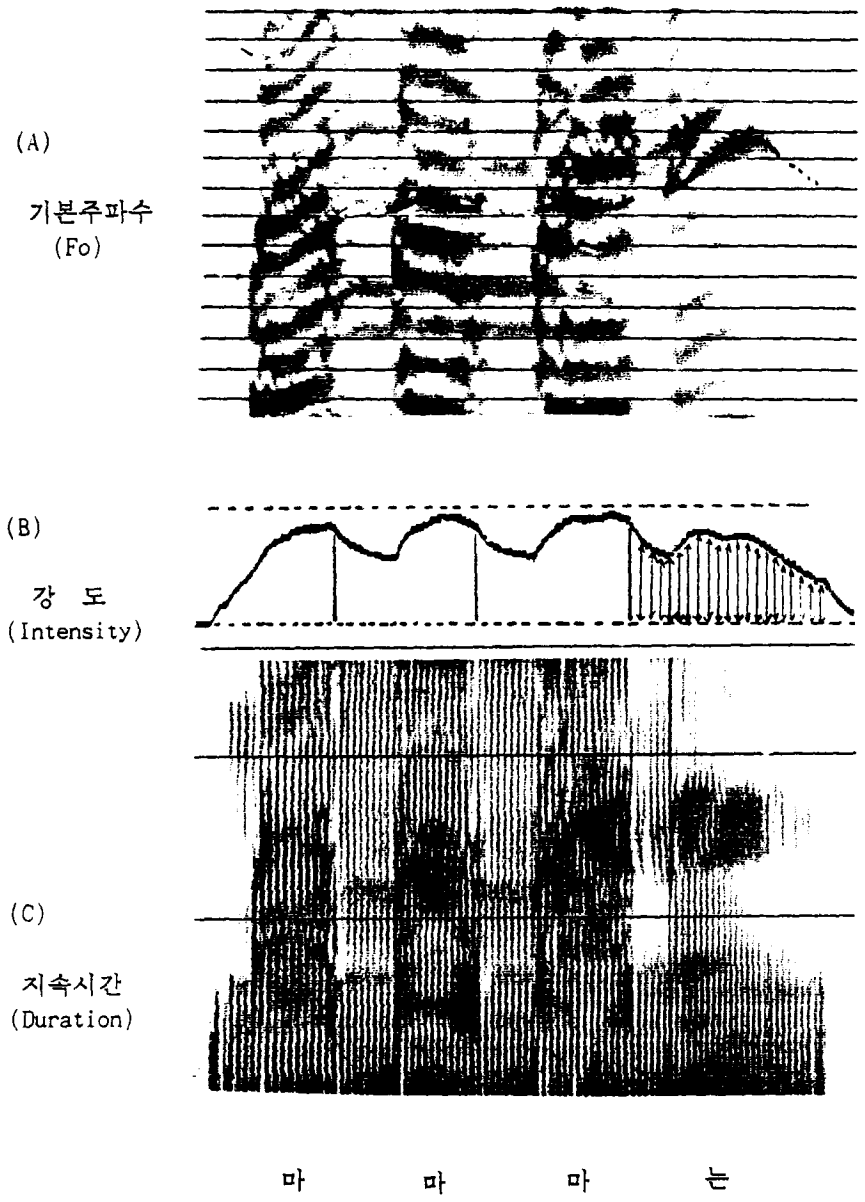


도표 2. 무의미 낱말 "마마마는"에 해당하는 세가지 음향 성분을 보여주는 스펙트로그램: (A) 고저(Pitch)
(B) 강도(Intensity), (C) 지속기간 (Duration)

Copyright (C) 2003 NuriMedia Co., Ltd.

(2) 강도와 지속시간

악센트 성분의 상관관계를 고려할 때 강도와 지속시간은 따로 분리하기가 어렵다. Beckman(1984)은 물리적인 지속시간과 세기의 관계는 밀접하게 상관관계가 있다고 보고하였으며, 역시 Pollack(1958)도 강도가 감소하면 지속시간도 감소한다고 보고하였다. 그러므로 세기의 지각은 에너지의 총계에 달려 있어서 스펙트럼상의 에너지 분포와 상관이 있고 지속시간도 말소리의 세기를 판단하는 데 중요한 실마리가 된다.

(3) 고저음과 지속시간

일반적으로 기계소리같이 일정한 정도의 음을 유지할 수 있는 순수음의 고저는 지속시간에 의해 영향을 받는다. 그러나 Beckman(1984)은 고저와 지속시간의 상호영향은 모음의 지속시간보다 훨씬 짧은 자극에 제한된다고 하며, 말소리의 지속시간은 고저음의 측정에 영향을 주지 못한다고 주장하였다. 그렇다면 고저음과 지속시간의 상관관계를 고려할때 순수음과 인간의 말소리에 차이가 있음을 인정하여야 할 것이다.

3.3. 표준한국어 악센트 특성

(1) 악센트 성분

가. 단순성분

악센트의 음향 성분 중에서 가장 두드러진 것을 한 가지만 정한다면 고저라 할 수 있다. 도표 3 에서 볼 수 있듯이 말소리를 대략 다섯 단계(아주 낮게, 낮게, 보통으로, 높게, 아주 높게)의 음역으로 나누어서 발음한 후 악센트 받는 음절에 해당하는 부분을 분석한 결과, 고저는 모든 피실험자 공히 음역이 올라감에 따라 일관성 있게 높아지고, 강도는 개인에 따라 다르지만 전체적으로 음역에 상관없이 불규칙하게 나타나며, 그리고 지속시간은 음역의 정도에 별 영향을 받지 않는다는 사실을 확인할 수 있다. 그러므로 표준한국어 악센트의 성분을 단순 성분론으로 기술한다면 가장 현저한 성분은 고저가 틀림없다. 그리고 강도와 지속시간 중 순위를 정한다면 지속시간이 먼저라고 할 수 있다. 악센트 받는 음절의 지속시간이 그렇지 않은 음절보다 낭독체에서 보통 30%에서 80% 그리고 자연스런 대화체에서 100%에서 300% 정도까지 길어지는 반면에, 강도는 개인차가 더 심하여 일관성이 약하기 때문이다.

지금까지 단순성분론을 주장해 온 예로는 이숭녕(1960)의 강약설과 정인섭(1965) 등의 고저설이다. 전자는 음향기구로 측정한 것이 아니고 다만 청취 판단에 따른 것으로, 본 실험 결과를 기준으로 해석하였을 때 받아들일 수 없다. 반면에 후자는 최초로 표준한국어 악센트를 스펙트로그램으로 분석한 실험음성학 연구 결과로 얻어진 것으로서, 본 연구와 실험방법에서는 다소 다르지만 고저설에서 일치한다고 볼 수 있다.

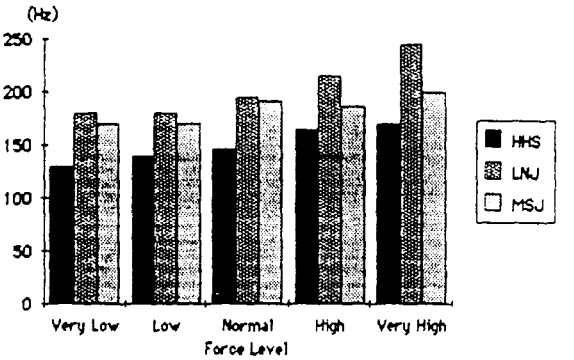
나. 복합성분

복합성분론은 악센트가 두가지 이상의 성분으로 구성된다는 것을 가정으로 성분간의 상관관계가 있음을 명시하는 것이다. 앞에서도 언급하였듯이 악센트는 세가지 음향 성분이 복합적으로 이루어진 것이 틀림없으나 문제가 되는 것은 성분간의 비율인데 그 비율을 정확히 측정하기란 대단히 어렵다. 그 이유로는 여러가지가 있겠으나 대략 다음과 같다. 개인마다 다른 발성기관, 발화방법 및 습관, 성별 차이, 연령 차이 등이라 할 수 있다.

Copyright (C) 2003 NuriMedia Co., Ltd.

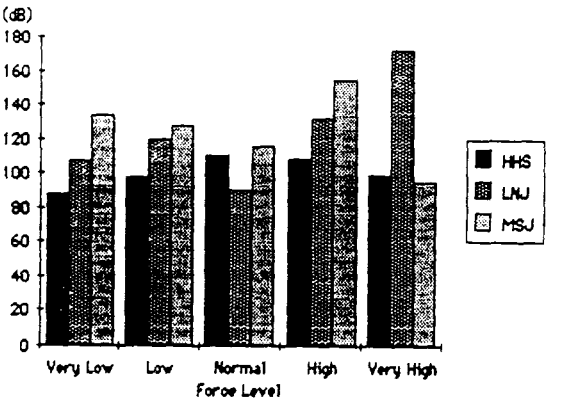
(A)

기본주파수
(Fo)



(B)

강 도
(Intensity)



(C)

지속시간
(Duration)

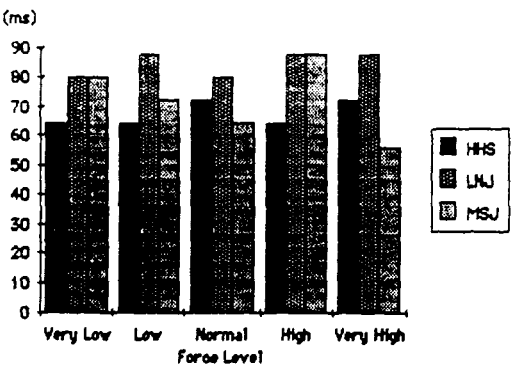


도표 3. 화자의 음역을 다섯 단계로 구분하여 발화한 "마마마는"에서 구절악센트 받는 "는"의 모음에 해당하는 음향 성분 비교: (A) 기본주파수, (B) 강도, (C) 지속시간

Copyright (C) 2003 NuriMedia Co., Ltd.

아직까지 이러한 문제를 해결하려는 연구는 이루어지지 않고 있다. 다만 단순성분론에서 기술한 것을 기초로 복합성분론을 논한다면, 표준한국어의 악센트는 고저, 지속시간 그리고 강도가 복합적으로 구성되었다고 할 수 있으며 구성 비율은 위에 나열한 순서라 할 수 있다. 이현복(1982)은 표준한국어 악센트가 강약, 장단 및 음질로 이루어진다고 하였다. 이러한 주장은 청취판단에 의한 것인데 실험음성학 결과로 얻어진 본 결과와 일치하지 않는다. 가장 큰 차이점은 고저인데 악센트 성분에서 고저 성분이 포함되지 않는 언어는 아직 보고되지 않았다는 점만 고려하여도 위의 주장이 옳지 못하다는 것이 분명해 진다. 강약은 개인차가 심하다는 것이 실험자료로 확인되었고, 음질의 차이는 절대적으로 개인차에 달려 있으므로 객관적 자료로 판단되어야 할 악센트 성분에 음질을 포함하는 것은 옳지 않다고 본다.

(2) 악센트의 위치

가. 어휘악센트

어휘악센트는 낱말 자체 내에서 일정한 음절에 인접 음절보다 특정한 성분이 현저하고 일관성있게 나타난다고 가정하자. 그렇다면 표준한국어의 악센트 성분이 고저로 확인되었기 때문에 어휘 악센트 위치는 고저를 기준으로 측정하는 것이 타당할 것이다. 2음절과 3음절로 구성된 무의미 낱말 '마마'와 '마마마'를 문장의 틀에 삽입하여 스펙트로그램상에서 악센트 위치를 확인하였다. 일반적으로 음성학실험에서 무의미 낱말을 사용하는 이유는 분절음의 음향성분 차이에서 오는 영향을 최소화하기 위한 것으로 특히 운율연구에서 주로 무의미 낱말을 사용한다.

2음절어는 어두와 어중에서 둘째 음절이 높고 어미에서는 첫째 음절이 더 높다. 그리고 3음절어에서는 모든 위치에서 둘째 음절이 첫째 음절이나 세째 음절보다 높다. 그러나 첫째 음절과 세째 음절간에 차이는 위치마다 다르다. 어두에서는 세째 음절이 높고, 어중에서는 거의 비슷하며, 어미에서는 첫째 음절이 더 높다. 스펙트로그램상의 기본주파수를 해독할 때에 애매모호한 점이 있다. 예를 들어 어중에서 2음절어와 3음절어의 첫째와 둘째 음절 사이가 곡선으로 이어지기 때문에 그 차이를 비교하기가 어렵다. 여기서는 보통 흔히 사용하는 음절에서 이어지는 선의 중간 지점을 따서 해석하였다(도표 4 참조).

나. 구절악센트

위에서 확인하였듯이 표준한국어의 어휘악센트는 음절 간의 차이가 미약하다. 어휘내의 음절 간의 우세가 문장 위치에 따라 변하므로 고정된 어휘악센트는 없다고 보아야 마땅할 것이다. 오히려 구나 절의 마지막 음절에 악센트가 뚜렷하게 나타남을 확인할 수 있는데 이러한 현상으로 미루어 보아 표준한국어는 구절악센트가 있다고 말할 수 있다(Koo, 1986).

도표 5-A 에서 볼 수 있듯이 문장에서 각 구의 끝 음절 '는', '를' 및 '해'에 고저로 나타나는 악센트가 온다. 도표 5-B에서도 절의 마지막 음절 '면'에 가장 높은 악센트가 왔다. 한국어의 구나 절의 끝 음절은 대부분 토씨들이기 때문에 악센트가 토씨에 온다고 말하기 쉬우나 토씨가 생략될 때는 낱말의 마지막 음절에 악센트가 온다. 그러므로 발화 단위나 억양구 끝 음절에 악센트가 온다고 말할 수 있다. 이러한 구절악센트는 억양곡선과 밀접한 관계가 있으므로 억양과 같이 연구되어야 한다.

Copyright (C) 2003 NuriMedia Co., Ltd.

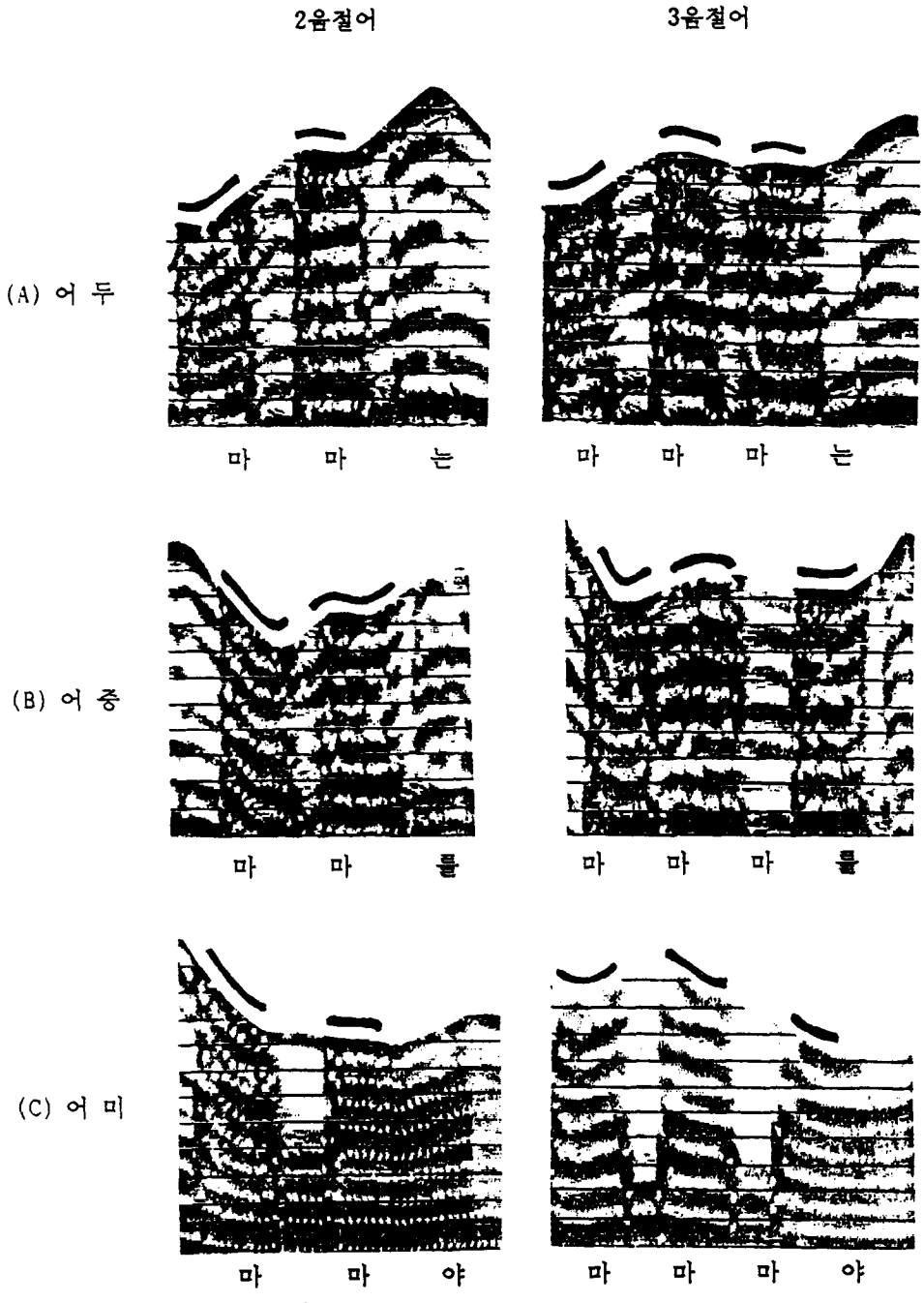
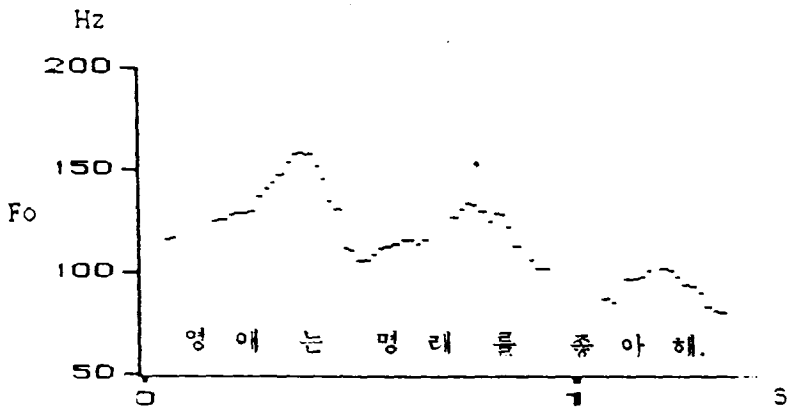


도표 4. 문장 위치에 따른 2음절어와 3음절어의 어휘악센트 위치

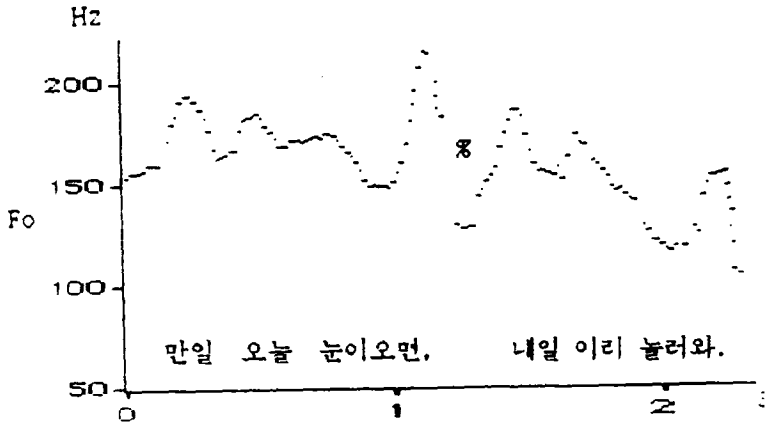
Copyright (C) 2003 NuriMedia Co., Ltd.

(A)



'영애는 명래를 좋아해'에서 각 구의 마지막 음절
'는', '를', '해'에 구절악센트가 온다.

(B)



'만일 오늘 눈이오면, 내일 이리 놀러와'에서
절의 마지막 음절 '면'에 구절악센트가 온다.

도표 5. 기본주파수 곡선상에 나타난 구절악센트

Copyright (C) 2003 NuriMedia Co., Ltd.

(A)

'명래는'에서 구절악센트에 강조악센트가 덧붙여져
'는'이 인접 음절보다 현저하게 더 높다.

(B)

'명래는'에서 두번째 음절 '래'에 구절악센트 받는
'는'보다 더 높은 강조악센트가 온다.

도표 6. 기본주파수 곡선에 나타난 두가지 강조약센트 형태

Copyright (C) 2003 NuriMedia Co., Ltd.

다. 강조악센트

보통 속도나 음역으로 말할때는 악센트가 구절의 마지막 음절에 오나 특정한 부분을 강조할 때는 두가지 현상으로 나타난다. 구절악센트에 덧붙여져 보통 구절악센트보다 끝음절이 훨씬 높게 나타나거나(도표 6-A), 아니면 낱말의 둘째 음절에 악센트가 와서 구절악센트보다 더 높게 나타난다(도표 6-B). 이러한 두 가지 패턴의 의미적 차이는 화용론적 입장에서 해석할 문제다.

4. 음성합성시 악센트정보 적용

지금까지 음성합성 과정에서 주로 운율처리에 관한 것을 중심으로 기술하고 악센트 특성을 분석하였다. 이러한 분석자료는 음성합성에 직접 적용하기에 아직 미흡하다. 운율에 대한 음향특성 파라미터를 설정하기 위해서는 해당 자료를 정확한 수치로 측정하고 통계처리를 하여야 한다. 운율처리에 유용한 자료를 얻기 위해서 우선 선행되어야 할 것은 운율의 특성을 충분히 이해하고 타당한 분석 방법을 도입하는 일이다. 운율 분석에서 해결하여야 할 사항을 몇가지 나열하면 다음과 같다.

(1) 음절 길이 조절

발화단위 내에 음절 수가 적을 때는 각 음절의 지속시간이 길어지고 역으로 음절 수가 많아질수록 각 음절의 지속시간은 짧아진다. 이와 같이 음절 수에 따라 길이가 조정되는 현상과 발성속도에 따른 정확한 길이 조절 규칙을 찾아내야 한다.

(2) 악센트 위치 변경

앞에서 한국어의 악센트를 어휘악센트, 구절악센트 및 강조악센트로 구분하여 분석한 결과 문장의 위치에 따라 악센트가 변하는 것을 확인하였다. 문장의 종류와 길이에 따라 악센트 위치가 어떻게 변하는가를 조사한 다음 그에 따른 법칙을 설정하여야 한다. 음운론적 규칙을 정하기 위해서는 가장 현저하게 나타나는 악센트 위치만 중시하였으나 음성합성에서는 악센트 받는 음절 뿐만 아니라 인접 음절의 성질도 참조하여야 한다. 예를 들어 3 음절어의 경우 각 음절간의 상대적인 높낮이, 강약 또는 음장의 차이도 충분히 활용할 수 있도록 하여야 한다.

(3) 운율 성분간의 상관 관계

음질은 고저, 강도 및 지속시간의 복합적인 상관관계에 의해 결정된다. 힘이 있어 보이는 말소리는 강도가 강한 것이고, 여성이나 어린이의 말처럼 높낮이가 심한 말은 고저의 차가 큰 것이며 그리고 속도가 느린 말은 보통 남성이나 어른들의 특징인데 비교적 지속시간이 길어지는 경향이 있다. 이처럼 화자마다 다른 발화 습관이 있으므로 연령 및 성별에 따라서도 충분한 자료를 구하여야 한다.

5. 결 론

지금까지 음성합성의 과정을 소개하면서 운율정보 추출과 운율합성에 역점을 두고 기술하고, 표준한국어 운율의 주요 자질인 악센트를 위주로 분석방법과 절차를 소개하였다. 또한 운율 분석 자료가 음성합성에 적용되기 위해서 해결되어야 할 중요한 문제점들을 열거하고 기술하였다. 그 중에서 근본적으로 해결되어야 할 사항으로 운율합성

Copyright (C) 2003 NuriMedia Co., Ltd.

율 위한 파라미터 데이터베이스에 저장될 자료는 명확한 규칙과 통계처리에 의한 것이어야 하며, 적절한 운율 분석 방법이 도입되어야 한다. 고차원적인 자연스런 연속 음성을 합성하기위해서 앞으로 수행되어야 할 운율연구 내용은 남녀 성별과 연령층을 고려한 충분한 수의 피실험자 선정, 실험실에서보다 실제 발화 현장에서 얻은 자료 분석, 풍부한 감정이 포함된 다양한 종류의 발화에서 생성되는 운율자료 수집, 그리고 속도에 따른 운율규칙 설정 등이라 할 수 있다.

참 고 문 헌

구희산 (1993). 음성인식과 합성을 위한 운율분석. 「HCI 93 학술대회 발표 논문집」 2.1. 93-100.

안승권, 성광모 (1992). 포만트 궤적 중첩방법에 의한 한국어 문자-음성 변환. 「음성통신 및 신호처리 워크샵 논문집」 9, 197-200.

이승녕 (1960). 현대 서울말의 악센트 연구. 「국어학논고」 179-223. 동양출판사

이양희 (1992). 음성합성기술 현황 및 전망. 「음성통신 및 신호처리 워크샵 논문집」 9, 88-96.

이현복 (1982). 한국어리듬의 음성학적 연구. 「말소리」 4. 31-48. 대한음성학회.

정인섭 (1965). 우리말 악센트는 고저(pitch)악센트다. 「중앙대학교 논문집」 10. 9-50.

최운천 외 4명 (1992). 고품질의 한국어 문장음성변환 시스템: 글소리 II. 「음성통신 및 신호처리 워크샵 논문집」 9, 193-196.

Beckman, M. (1984). Toward Phonetic Criteria for a Typology of Lexical Accent in Japanese. Diss. Cornell Univ.

Koo, H-S. (1986). An Experimental Acoustic Study of the Phonetics of Intonation in Standard Korean. Diss. Univ. of Texas at Austin. [Hanshin Publishing Co.].

Pollack, I. (1958). Loudness of Periodically Interrupted White Noise. *JASA* 30, 181-85.

Snow, W. (1936). Change of Pitch with Loudness at Low Frequencies. *JASA* 8, 14-19.

Stevens, S. (1935). The Relation of Pitch to Intensity. *JASA* 6, 150-540.

Stevens, S and J. Volkman (1940). The Relation of Pitch to Frequency: a revised scale. *Journal of Psychology* 53, 329-53.

Terhardt, E. (1975). Influence of Intensity on the Pitch of Complex Tones. *Acoustica* 33, 345-48.

서울시 동작구 흑석동 221
중앙대학교 영어교육학과
156-756
Fax: (02) 816-9938

Copyright (C) 2003 NuriMedia Co., Ltd.