자연스러운 여성 합성음을 위한 한국어의 피치 변화 법칙

The Rule of Korean Pitch Variation for a Natural Synthetic Female Voice

김 정 원*, 박 대 덕*, 김 보 현*, 권 철 홍*
(Chung Won Kim*, Dae Duck Park*, Boh Hyun Kim*, Cheol Hong Kwon*)

요 약

본 논문은 자연스러운 여성 합성음을 위한 피치 변화 법칙을 세웠다. 피치 변화 법칙이 적용되는 기본 단위, 즉 억양구는 주로 어절(들)로 이것의 첫번째, 두번째, 마지막 음절의 피치값을 연결해 피치 변화 곡선을 형성하였는데, 첫번째, 두번째 음절의 피치값은 각 음절의 초성에 따라, 마지막 음절의 피치값은 기능어의 종류에 따라 결정되었다. 억양구 사이에는 '쉼(pause)이 있는 경계' 또는 '쉼이 없는 경계'가 오며, 쉼이 있는 경계에는 relaxation이 있다. 이러한 억양구의 피치 변화 곡선, 경계 현상들이 모여 한 문장의 피치 패턴을 만들었다.

ABSTRACT

In this paper we make a rule of pitch variation for a natural synthetic female voice. Intonation phrase, which is the basic unit the rule is applied to, mostly consists of a syllable or syllables. The pitch values of the first, second, and final syllables make up the pitch contour of the intonation phrase. Those of the first and second syllable are determined by the initial consonants of the respective syllables, and that of the final syllable by the type of the function word. There are two kinds of boundaries between intonation phrases. One is a boundary with pause, and the other is a boundary without pause. The pitch contour of the intonation phrase with the boundary phenomena determines the pitch pattern of a sentence.

Ⅰ. 머리말

요즈음 음성합성에 대한 관심이 커지면서 학교나 연구소, 기업등에서 음성합성에 대한 연구가 활발히 진행되고 있다. 음성합성의 초기에는 합성음이 인간에게 정확하고 바르게 전달되는가, 즉 인간이 합성음을 듣고 바르게 이해할 수 있는가 하는 음의 명료성(intelligibility)에 중점을 두었다. 그러나 의미 전달은 가능하다 할지라도 이것은 단순한 기계음에 불과하다. 합성음의 궁극적인 목표가 인간 언어라면 명료성외에 자연성(naturalness)이 고려돼야 한다. 자연성은 운율(prosody)과 관계가 있으며, 운율은 지속시간(duration), 세기(intensity), 피치(pitch)의 3가지 요소가 어우러져 이루어진다. 우리말의 경우, 지속시간이 가장 큰 비중을 차지하나[6], 3가지 운율 요소 중 어느 하나라도 간과하면 자연스런 한국어와는 거리가 있다. 자연성에 가장 큰 영향을 미치는 지속시간에 대해서는 많은 연구가 행해졌으나 피치, 세기에 대해서는 체계적이고 광범위한 연구가 부족한 실정이다. 그러나 지속시간 못지 않게 피치도 상당히 중요하다. 피치가 없허지지 않을 경우 발화는 평이하고 단조롭게 들리며, 잘못된 피치가 얹혀질 경우, 무대위의 배우가 책을 읽듯이 대사를 말하는 것처럼, 어색하게 들린다. 피치 패턴은 상황과 문맥에 따라 다양하게 변한다. 합성음에 대화체의 생생한 현장감 또는 낭독체의 차분함을 줄 수 있는 것은 피치다.

본사는 개발중인 여성 목소리의 음성합성 프로그램 가라사대 II가 기존 가라사대 I의 무미건조하고 기계적인 소리를 탈피하고 좀 더 밝고 생기있는 소리로 전환하는 것을 목표 중 하나로 삼았으며, 이를 위해 자연스럽고 생생한 어투에 적합한-낭독체 보다는 대화체적 성격이 더 강한-여성용 피치 규칙이 필요하였다. 이에 본 논문은 실험을 통해 밝고 생기있는 여성 발화의 피치 경향을 분석하고, 규칙을 도출하여, 가라사대 II에 실제로 적용하는 것을 목표로 삼았다. 이전 연구들은 주로 낭독체에서 문장 이상의 단위에 나타나는 개략적이며 전체적인 피치 현상에 관한 것이었다. 이러한 피치 규칙은 발화 전체적으로 나타나는 기본적이며 규칙적인 피치 패턴을 추출한 것으로 대부분의 발화에 두루 적용되는 보편성을 지닌다. 따라서 듣기에 안정적이나 어느 문장이나 그 피치 패

*디지콤 정보통신연구소
접수일자 : 1996년 7월 11일

턴이 바슷하기 때문에 자칫 단조로울 수 있다. 또한 실제 발화에서 문장마다 다르게 나타나는 다양한 피치 패턴을 반영할 수 없다. 자연스런 생동감을 주기위해서는 개략적이며 전체적인 피치 현상 외에 분절음의 영향과 같이 미세한 피치 현상에 관한 규칙이 필요하다. 이러한 피치 규칙은 피치가 국부적으로 실현돼 세부적인 피치값까지 얻을 수 있어 발화에 악센트를 주며 생생한 리듬감을 줄 수 있다. 이상적인 피치 규칙은 전체적인 피치 현상과 미세한 피치 현상이 적절히 조화될 때 가능한데, 본 연구가 특히 발화의 생생함을 살리는데 중점을 둔 만큼 미세한 피치 현상을 위주로 하여 피치 규칙을 수립하고, 전체적인 피치현상을 보완하는 방식을 취하였다.

Ⅱ. 실험 과정

2.1 실험 목표

본 논문에서 중점적으로 살펴 볼 사항은 다음과 같다.

1. 하나의 피치 변화 곡선이 형성되는 단위는 무엇인가?
2. 문장 전반에 걸쳐 피치 baseline은 어떻게 변하는가?
3. 피치 변화 곡선을 결정하는 요소는 무엇인가?
4. 경계가 있을 때 어떠한 피치 변화가 있는가?
5. micro melody를 실현하는 요소는 무엇인가?
6. 조사에 따른 피치 변화는?
7. 문장의 종결형 어미에 따른 피치 변화는?

2.1. 피실험자

서울 태생으로 조음 훈련이 잘 된 서울대 언어학과 음성학 박사과정의 30대 초반 여성을 피실험자로 삼았다.

2.2. 실험 자료

1. 실험 목표에 제시된 전반적인 사항들을 살펴 보기 위해 국어 문법서에서 너무 짧거나 단순하지 않은, 절(clause)을 포함하는 문장들을 선택하였다. 이러한 문장들은, 특히 실험 목표 4의 경계 현상을 살펴 보기 위해, 한 번은 절 경계를 띄어 읽고 다른 한 번은 절 경계를 붙여 읽도록 피실험자에게 지시하였다. 실험 문장은 54개이나, 이처럼 두가지 방식으로 발화하였으므로 전체로는 108개의 문장이 사용된 것과 같다. 실험 문장 가운데 일부를 표 1에 나타냈으며, 절 경계를 '/'로 표시하여 이 부분을 붙여 읽고 띄어 읽게 하였다.

표 1. 실험 문장

실험 문장
구름 한 조각이 사라지듯/사람의 목숨도 사라진다.
열심히 노력하는 학생이/성적이 좋다.
그이가/"내일 가거라"/라고 말했다.
호랑이는 죽어서 가죽을 남기고/사람은 죽어서 이름을 남긴다.
밖이 몹시 추우니까/옷을 두껍게 입어라.
너는 이번 휴일에 네 방을 치우거나/차를 닦아야 한다.
비가 억수같이 퍼붓다가/갑자기 날씨가 화창해졌다.
네가 그렇게 생각했거든/하물며 남이야 말할 것 있나?
그는 또 집을 나갔는데,/이것은 그의 세번째 가출이었다.
소련의 개혁정책은,/내가 알기로는,/너무나도 많은 어려움에 직면해 있다.

위의 실험 문장들은 실험 목표 6에서 살펴 보고자 하는 조사 유형들도 포함하고 있다. 따라서 108개의 문장에 포함된 모든 조사들의 피치값을 측정하고, 이 값들을 조사 유형에 따라 분류하여 각 조사별로 평균값을 취하였다.

2. 실험 목표 5를 좀 더 자세히 살펴 보기 위하여 다음과 같은 실험 문장들을 만들었다.

(1) 여기에 가가가 있습니다.
(2) 여기에 기기기 있습니다.
(3) 여기에 귀귀귀 있습니다.
(4) 여기에 갸갸갸 있습니다.
(5) 여기에 각까까 있습니다.
(6) 여기에 간가가 있습니다.
(7) 여기에 갈가가 있습니다.
(8) 여기에 가까까 있습니다.
(9) 여기에 가카카 있습니다.
(10) 여기에 가사사 있습니다.
(11) 여기에 가싸싸 있습니다.
(12) 여기에 가마마 있습니다.
(13) 여기에 가하하 있습니다.
(14) 여기에 가라라 있습니다.

이러한 문장들을 기본 유형으로 하여, 밑줄친 어절의 첫째 음절의 초성 자음 'ㄱ' 에 18개의 자음을 대치하여 각각 14개의 동일한 문장들을 만들었다. 총 266(19 * 14)개의 문장에서 주변 음소에 따른 모음의 피치 변화를 살펴 보았다.

3. 1의 자료만으로는 실험 목표 7의 결과를 도출하기에 불충분하여 '여기에 책상이 있습니다.' 라는 문장을 기본 유형으로 하여, 평서문, 예/아니오 의문문, 의문사 의문문, 명령문, 청유문의 종결형 어미에 따라 35개의 문장들을 만들었다. 단 의문사 의문문은 '여기에 무엇이 있습니까?' 를 기본 문장으로 하였다.

2.3. 녹음 방법

사설 스튜디오의 녹음실에서 409개의 문장을 하루에 1번씩 3일에 걸쳐 3회 발화시켰다.

2.4. 측정 방법

테이프에 녹음된 실험자료를 A/D Converting 하여 PC에 저장한 뒤 CSL(Computerized Speech Lab)[15]을 이용해 피치값을 측정하였다. 피치는 모음의 안정구간에서

측정하는 것을 원칙으로 하였다.

Ⅲ. 실험 결과

실험 결과에 의하면 피치가 얹히는 단위, 즉 억양구(intonation phrase)는 경계와 경계사이의 어절(들)이다. 이러한 억양구는 주로 어절인데, 어절 이상의 단위일 수도 있다. 예를 들면 "할 수 있다."는 3개의 어절이 하나의 억양구를 이룬다. 억양구가 소단위인 것은 실험 문장이 대체적으로 짧고 고립된 문장이라는 점과 피실험자의 발화 양식이 밝고 생기있다는 점에 어느 정도 영향을 받은 듯 하다. 피치 변화 곡선은 억양구 단위로 재조정되며, 문장이 진행됨에 따라 피치가 점차 하강한다는 문장 전반에 걸친 하강(declination)현상은 발견되지 않았다. 오히려 문장 전체에 걸쳐 평평한 피치 baseline이 생겼는데, 이것은 피실험자의 평균 기본주파수 260-270(Hz) 사이에 존재한다. 이러한 결과는 "대화문에서 문장내 많은 곳에 뾰족한 peak가 두드러지지만 문장 전체적으로는 평평한 baseline이 생긴다"는 고 도홍[1]의 실험 결과와 일치한다.

억양구 단위로 얹히는 개별적인 피치 변화 곡선은 억양구를 이루는 음절들의 피치값에 따라 결정되었다. 1음절로 된 억양구에서는 그 음절의 피치값이, 2음절로 된 억양구에서는 그 두 음절의 피치값들이 피치 변화 곡선을 결정하였다. 3음절 이상으로 된 억양구의 경우 그 어절의 첫 음절, 둘째 음절, 마지막 음절의 피치값이 결정하였다. 시작 피치값, 마지막 피치값 외에 중간 피치값으로 둘째 음절의 피치값을 가졌는데, 한국어의 어절(들)(억양구) 구조가 주로 의미어+기능어의 결합이라는 점을 고려할 때 첫번째, 두번째 음절의 피치값은 의미어의 피치 패턴을 결정하며 마지막 음절의 피치값은 기능어의 피치 패턴을 결정한다. 이러한 결과는 표준 한국어의 낱말 악센트가 2, 3, 4음절 낱말을 대상으로 할 때 그 위치가 처음 두 음절로 고정된다[10]는 법칙과 무관하지 않다.

이처럼 억양구를 이루는 음절들의 피치값에 따라 결정된 피치 변화 곡선은 그 뒤에 두가지 유형의 경계가 온다. 하나는 대부분의 억양구 뒤에 오는 경계로 쉼(pause)이 없다. 그리고 쉼이 없는 대신 경계임을 나타내기 위해 운율상으로 그 경계 앞 억양구의 마지막 음절이 길어지고 대체적으로 억양구의 끝이 상승하는 피치 변화가 있다. 이러한 피치 변화는 한국어 억양은 구단위에서 오름조 경향을 띤다는 구 회산[3]의 주장과 일치한다. 이것을 '쉼이 없는 경계'라고 부르기로 한다. 다른 경계는, 운율 변화는 쉼이 없는 경계에서와 같다. 즉 경계 앞 억양구의 마지막 음절이 길어지고 억양구의 끝이 상승하는 피치 변화가 있다. 차이점으로 이 경계에는 쉼이 있고 마지막 음절의 끝부분에서 그 음절의 피치값이 130-140Hz까지 떨어지는(놀어지는)현상이 발생한다. 이처럼 피치값이 경계앞에서 놀어지는 현상을 relaxation이라 하고, relaxation이 생기는 경계를 '쉼이 있는 경계'로 부르기로 한다. 따라서 쉼이 있는 경계는 쉼이 없는 경계에 비해 보다 그 구분이 명확하며, 발화 속도에 관계없이 반드시 띄어 읽어야 하는 절의 뒤 등에서 볼 수 있다.

그러면 각 음절의 피치값을 결정하는 요인은 무엇일까? 본 실험에서는 분절음의 영향이 상당히 큰 것으로 나타났다. 모음의 피치값은 선행하는 초성 자음에 의해 그 값이 결정되었다. 따라서 그 음절의 초성이 19개의 자음 중 무엇이냐에 따라 음절의 피치값이 달라진다. 실험 결과 앞 초성에 따른 모음의 피치값은 표 2와 같다.

표 2. 앞 초성에 따른 모음의 피치값(분석에 의한 피치값)
(단위:Hz)

초성	모음의 피치값	초성	모음의 피치값	초성	모음의 피치값	초성	모음의 피치값	초성	모음의 피치값
ㄱ	236	ㅁ	229	ㄲ	300	ㅋ	329	ㅅ	322
ㄷ	246	ㄴ	229	ㄸ	334	ㅌ	345	ㅎ	345
ㅂ	238	ㅇ	231	ㅃ	315	ㅍ	345		
ㅈ	246	ㄹ	230	ㅆ	319	ㅊ	345		
				ㅉ	344				

표 2에서 볼 수 있는 것 처럼, 평균 기본 주파수 260-270Hz와 비교해, 마찰음을 제외한 평음(ㄱ, ㄷ, ㅂ, ㅈ), 유성 자음(ㅁ, ㄴ, ㅇ, ㄹ)뒤의 모음이 비슷하게 낮은 피치값을 갖고, 경음, 격음, 마찰음 뒤 모음의 피치값이 비슷하게 높다. 이것은 구 회산[3], Umeda[14]의 무성 자음 뒤 모음의 기본 주파수는 높고 유성 자음 뒤 모음의 기본 주파수는 더 낮다는 주장과 어느 정도 일치한다.

마지막 음절의 피치값은 앞 초성에 의한 영향보다, 그 음절와 기능어의 종류에 따라 영향을 받았다. 문장 내에서 마지막 어절이 아닌 경우, 조사의 종류에 따라 마지막 음절의 피치값이 결정되었고, 마지막 어절인 경우, 문장 종결 어미의 유형에 따라 그 피치값이 결정되었다. 문장의 마지막 어절이 아닐 때 조사별 피치값은 표 3과 같다.

표 3. 조사의 종류에 따른 피치값(단위:Hz)

조사	예	피치값
주격 조사	-이, -가	300-310
목적격 조사	-을, -를	280-290
인용격 조사	-라고	300-310
비교격 조사	-보다	290-300
대조, 강조 조사	-은, -는, -만, -도, -조차, -까지	340-350

표 3에 속하지 않는 다른 조사들의 피치값은 첫 음절, 둘째 음절 피치값에서 처럼 선행하는 초성자음의 종류에 따라 결정되었다.

문장의 마지막 어절에서 문장 종결 어미 유형에 따른

표 4. 문장 종결 어미 유형에 따른 피치값(단위 : Hz)

문장유형	문장 종결형 어미	피치값
평서문	-습니다	152
	-다	154
	-죠	154
	-소	150
	-어	152
	-구나	158
	-어요	152
	-네	* 273 + 135
예/아니오 의문문	-습니까	385
	-느냐	334
	-니	345
	-지	370
	-어	344
	-소	417
	-죠	386
의문사 의문문	-습니까	146
	-느냐	154
	-니	143
	-지	* 343
	-어	145
	-소	151
	-죠	* 357
명령문	-어라	161
	-어	156
	-게	147
	-려무나	150
	-렴	161
	-구료	149
	-소서	145
청유문	-(ㅂ)시다	143
	-자	149
	-세	134

피치값은 표 4와 같다.

평서문, 명령문, 청유문은 종결형 어미의 피치값에서 볼 수 있듯이 주로 하강조(fall)이다. 의문문의 경우 예/아니오 의문문은 상승조(rise), 의문사 의문문은 주로 하강조이다. 예외적으로 평서문의 종결형 어미 '-네'는 피치값이 상승했다 떨어지는 상승-하강조(rise-fall)이며, 의문사 의문문 중 종결형 어미 '-지, -죠'는 상승조이다.

Ⅳ. 피치 변화 법칙

실험 결과를 기초로 가라사대 II에 적용할 피치 규칙을 세웠다. 본 규칙은 글말(written language)이 아니라 소리말(spoken language)에 적용되는 규칙이다. 다시말하면 입력된 텍스트가 소리나는 형태로 전환된 후(음운 규칙을 거친 후)에 이러한 규칙이 적용된다.

실제 발화에서 얻은 피치값을 수정하지 않고 합성음에 그대로 적용하는 데에는 무리가 있었다. 문장 합성 결과 지나치게 높거나 낮은 피치값이 얻혀질 경우 그 음이 깨지거나 우는 소리가 섞였다. 더 구체적으로 말하면, 사람이 지나치게 높은 음을 낼 때 소리가 가늘어지고 음이 갈라지며, 낮은 음을 낼 때 소리가 굵어지고 '웅웅'거리는 현상이 합성시 나타났다. 이처럼 실제 발화에서 다양하게 큰 폭으로 변하는 피치값들을 합성음이 소화해 내기 어려웠다. 그래서 높은 피치값들은 많은 문장을 합성해 본 후 점차적으로 그 값을 낮췄으며, 낮은 피치값들은 그 값을 높혔다. 본 장에서 언급될 피치값들은 이러한 과정을 거친 최종값들이다. 하나의 억양구 내에서 최대 피치값과 최소 피치값의 차가 클 경우에도 비슷한 현상이 나타나므로 피치 변이폭을 제한하여 최대값과 최소값을 조정하였다. 또한 자연음에서 조사와 종결 어미의 종류에 따라 다양하게 변하는 피치값들을 그 값 그대로 합성음에 적용할 수도 있었으나, 조사와 어미 간의 피치 변이폭이 그리 크지 않으므로 규칙을 좀 더 간결하게 하기 위해 합성음의 자연성에 영향을 주지 않는 범위 내에서 그 값들을 평준화, 단순화시켰다.

다음에 기술될 피치 규칙은 3장의 실험 결과로 부터 도출된 규칙이다. 기본적인 골격은 같으나 실험 결과의 피치값들이 합성음에 맞게 일부 수정되었다. 가라사대 II에 적용된 피치 규칙은 다음과 같다.

1. 피치 규칙이 적용되는 단위, 즉 억양구는 경계와 경계 사이의 어절(들)이다. 여기에 첫 음절, 둘째 음절, 마지막 음절의 피치값을 주고, 쉼이 있는 경계 앞에서는 relaxation값을 주어 그 값들을 연결한다.

2. 피치 규칙에 의해 변경된 앞 초성에 따른 모음의 피치값은 표 5와 같다.

표 5. 앞 초성에 따른 모음의 피치값(규칙에 의한 피치값)
(단위 : Hz)

초성	모음의 피치값	초성	모음의 피치값	초성	모음의 피치값	초성	모음의 피치값	초성	모음의 피치값
ㄱ	226	ㅁ	219	ㄲ	260	ㅋ	270	ㅅ	270
ㄷ	236	ㄴ	219	ㄸ	270	ㅌ	270	ㅎ	270
ㅂ	228	ㅇ	221	ㅃ	265	ㅍ	270		
ㅈ	236	ㄹ	220	ㅆ	269	ㅊ	270		
				ㅉ	270				

첫 음절과 두번째 음절에는 이 값이 그대로 적용된다. 그런데 고유 피치값이 서로 비슷한 음절들로 이루어진 억양구에서 (ex. 마니, 모여...) 피치 패턴의 단조로움을

피하고 끝이 상승 패턴을 이루도록, 마지막 음절에는 고유 피치값 + 20Hz를 준다.

3. 문장의 마지막 어절이 아닌 경우, 억양구의 마지막 조사가 아래에 열거한 조사 가운데 속할 때, 마지막 음절에는 고유 피치값 + 20Hz를 주지 않고, 해당 조사의 피치값을 준다. 피치 규칙에 의해 변경된 조사의 피치값은 다음과 같다.

(1)주격 조사 (-이, -가...) (280Hz)
예외)서술절 주어의 주격 조사에는 주격 조사의 피치값 280Hz를 주지 않고, 고유 피치값 + 20Hz (또는 30Hz)를 준다.
ex. 열씨미 노려카는 학쌩이 성저기 조타.:'성저기'는 서술절의 주어이므로, 마지막 음절의 피치값은 246Hz이다.
(2)목적격 조사 (-을, -를...) (280Hz)
(3)인용격 조사 (-라고...) (280Hz)
(4)비교격 조사 (-보다...) (280Hz)
(5)대조, 강조를 나타내는 보조사 (-은, -는, -만, -도, -조차, -까지...) (280Hz)
(6)그 외의 조사 (고유 피치값 + 20Hz)

4. 문장의 마지막 어절에서, 억양구의 마지막 음절의 피치값은 문장 유형과 문장 종결 어미 유형에 따라 결정된다. 피치 규칙에 의해 변경된 문장 종결 어미의 피치값은 다음과 같다.

(1)평서문 (190Hz)
주로 하강조(falling)의 피치 패턴
예외)종결어미가 '-네'일 때 피치값은 273Hz까지 올라갔다가 135Hz로 떨어진다. 즉, 피치값이 상승한 뒤 relaxation이 생긴다.
(2)예/아니오 의문문 (280Hz)
상승조(rising)의 피치 패턴
(3)의문사 의문문 (190Hz)
주로 하강조의 피치 패턴
예외)종결 어미가 '-지', '-죠'일 때 피치값은 290Hz로 상승조의 피치 패턴
(4)명령문 (190Hz)
하강조의 피치 패턴
(5)청유문 (190Hz)
하강조의 피치 패턴

5. 쉼이 있는 경계 앞에 relaxation이 있다. 피치 규칙에 의해 변경된 relaxation값은 170Hz로 일정하며 마지막 음절의 피치값과 연결한다.

V. 피치 규칙의 평가

위의 피치 규칙을 합성음에 적용한 뒤, 기존의 합성음과 비교해 자연성이 어느정도 향상되었는가를 살펴 보자. 또한 음성합성의 최종 목표는 자연음이므로, 합성음의 피치 패턴이 자연음의 피치패턴에 어느정도 근접했는지를 살펴보는 것도 필요하다.

자연성의 평가는 청취 실험을 통해 이루어졌다. 10명의 피실험자에게 가라사대 I과 본 피치규칙을 적용한 가라사대 II의 문장들을 들려 주고 자연성을 비교 평가하도

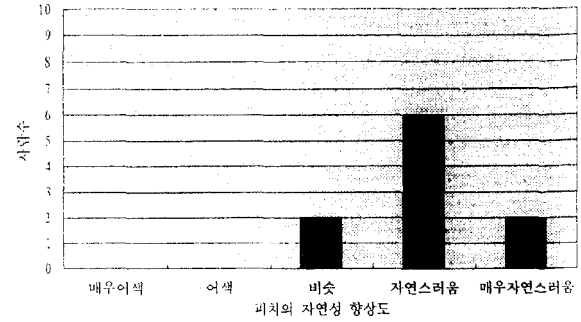
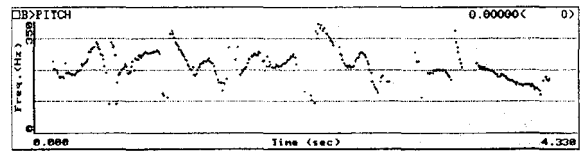


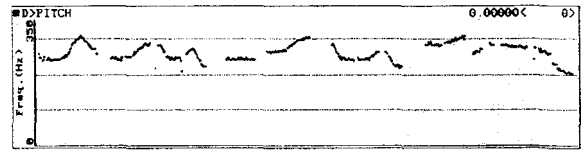
그림 1. 피치의 자연성 향상도

"너는 이 번 휴일에 내 방을 치우거나 차를 닦아야 한다.
(너는 이 번 휴이레 내 방을 치우거나 차를 다까야 한다.)"

<자연음>

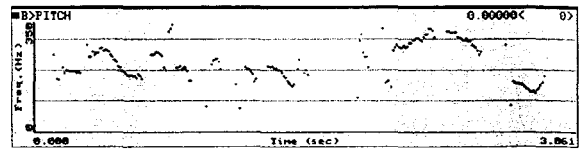


<합성음>



"내가 역에 도착하자, 기차는 떠났다.
(내가 여게 도차카자 기차는 떠날따.)"

<자연음>



<합성음>

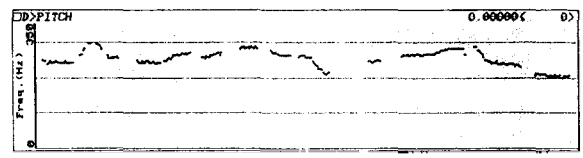


그림 2. 자연음과 합성음의 피치패턴 비교

록 하였다. 평가 단계는 가라사대 II가 가라사대 I과 비교해 1) 매우 어색하다. 2) 어색하다. 3) 비슷하다. 4) 자연스럽다. 5) 매우 자연스럽다. 와 같이 5단계로 구분하였다.

청취 실험 결과는 그림 1과 같다.

그림 1에서 보면 10명의 피실험자 가운데 2명이 '비슷하다', 6명이 '자연스럽다', 2명이 '매우 자연스럽다'에 응답하여, 최고점을 5로 할 때 전체 평균 4.0이다. 이것은 본 피치 규칙이 가라사대 II에 적용됨으로써 기존 가라사대 I에 비해 합성음의 자연성이 향상되었다는 것을 나타낸다.

합성음의 피치 패턴과 자연음의 피치 패턴을 비교해 봄으로써 피치 규칙의 일반성, 보편성을 평가 할 수 있다. 자연음으로 된 실험 자료의 문장들을 합성하여, 동일한 문장에 대해 합성음과 자연음의 피치 패턴을 비교한 결과는 그림 2와 같다. 피치 패턴은 CSL을 통해 출력하였다.

Ⅵ. 맺음말

실제 발화를 분석하여 음성합성 프로그램 가라사대 II에 적용할 피치 규칙을 세웠다. 본 규칙은 억양구를 단위로 하나의 피치 변화 곡선을 형성하였고, 이러한 피치 변화 곡선들이 모여 한 문장의 피치 패턴을 만들었다. 피치 변화 곡선은 억양구의 첫 음절, 둘째 음절, 마지막 음절의 피치값들을 연결하여 구현했고, 첫 음절, 둘째 음절의 피치값은 해당 음절의 초성 자음의 종류에 따라, 마지막 음절의 피치값은 문장의 마지막 어절이 아닌 경우에는 조사의 종류에 따라, 마지막 어절에서는 문장 유형과 문장 종결 어미 유형에 따라 결정되었다. 억양구 뒤에는 두가지 유형의 경계, 즉 쉼이 있는 경계와 쉼이 없는 경계가 왔다. 두 경계 모두에서, 경계 앞 억양구의 마지막 음절이 길어지고 주로 억양구의 끝이 상승하는 피치 변화가 있었으며, 특히 쉼이 있는 경계에는 피치값이 떨어지는(늘어지는) relaxation 현상이 생겼다.

자연스럽고 생생한 어투에 적합한 피치 규칙을 만드는 것이 그 목표였으며, 그러기 위해선 미세한 피치 현상이 실현되야 했다. 본 실험에서는 해당 음절의 초성 자음이 피치값에 미치는 영향을 발견함으로써 합성된 문장에 미세한 피치 현상을 줄 수 있었다. 그런데 이밖에도 여러가지 요인들이 피치값에 영향을 줄 수 있을 것이다. 그 첫번째로 모음의 영향이 있다. 구희산[3], Umeda[14]는 모음은 고유의 기본 주파수를 가지며, 고모음은 기본 주파수가 높고 저모음은 낮다고 하였다. 이처럼 모음이 세분화 돼 고모음, 중모음, 저모음에 따라 또는 전설 모음, 중설 모음, 후설 모음에 따라 피치값이 달라질 수 있다. 또는 21개의 모음 각각이 서로 다른 피치값을 가질 수도 있다. 자음의 영향으로는 본 규칙에 적용한 음절 내 초성 자음의 영향외에 중성 자음의 영향, 인접한 후행 음절의 초성 자음의 영향을 들 수 있다. 음절 내 초성 자음의 영향이 순방향적이라면 나머지 영향들은 역방향적이다. 분절음 외에 억양구 내에서 음절 위치도 피치값에 영향을 줄 수 있다. 환경이 되는 분절음이 같고, 동일한 모음이라도 억양구의 첫 음절에 위치하느냐, 중간 음절들에 위치하느냐, 마지막 음절에 위치하느냐에 따라 피치값이 달라질 수 있다. 이러한 변화 요인들을 복합적으로 고려한다면, 보다 정밀하며 환경에 따라 다양하게 변하는 피치값을 할 수 있을 것이며, 따라서 피치 규칙도 좀 더 자연음 규칙에 맞게 다듬어지고 보완될 것이다. 이것은 앞으로의 연구과제로 남겨두겠다.

본 논문은 피치 규칙에 인접한 초성 자음의 영향과 같은 미세한 피치 현상을 도입함으로써 합성음에 자연스런 생동감을 주었으며, 전체적이며 개략적인 피치 현상만으로는 어려웠던 합성음에 실제발화의 다양한 피치 변화 곡선이 구현되는 것을 가능하게 만들었다.

참 고 문 헌

1. 고도흥, Declaritive Intonation In Korean:An Acoustical Study of F0 declination, 한신출판사, 1988.
2. 구희산, An Experimental Acoustic Study of the Phonetics of Intonation in Korea, 한신 출판사, 1988.
3. 구희산, "억양패턴 분석방법", 제9회 음성통신 및 신호처리 워크샵 논문집, pp. 51-55. 1992.
4. 김민자, A Phonetic Study of the Structure of Fundamental Frequency Contours, 서울대학교 박사학위 논문, 1992.
5. 김상훈, 지민제, 최운천, "한국어 문장/음성 변환에서의 TD-PSOLA 적용", 제10회 음성통신 및 신호처리 워크샵 논문집, pp. 291-294, 1993.
6. 김연준, 오영환, "한국어 문서-음성 변환 시스템에서의 구문분석에 의한 운율조절에 관한 연구", 제10회 음성통신 및 신호처리 워크샵, pp. 285-290, 1993.
7. 김진영, 성굉모, "한국어의 억양에 관한 연구", Korea-Japan Joint Symposium on Acoutics, pp. 292-297, 1991.
8. 성철재, 한국어 리듬의 실험음성학적 연구-시간 구조와 관련하여, 서울대학교 박사학위 논문, 1995.
9. 이정철, 김상훈, "최소자승오차를 이용한 세그먼트 피치패턴 정형화", 제11회 음성통신 및 신호처리 워크샵 논문집, pp. 107-110, 1994.
10. 이현복, "현대 한국어의 악센트", 서울대학교 문리대 학보 19권 합병호(통권 28호), pp. 113-128, 1973.
11. 이호영, The Structure of Korean Prosody, 한신 출판사, 1990.
12. 임운천, 안수길, "독립 단어에서의 운률에 관한 연구", Korea-Japan Joint Symposium on Acoutics, pp. 280-291, 1991.
13. 지민제, 이용주, "한국어 Pause Pattern의 음향음성학적 분석", 통신처리를 위한 음성정보 변환기술 개발, 한국전자통신 연구소, 1990.
14. Noriko Umeda "Influence of segmental factors on fundamental frequency in fluent speech", Journal of Acoustical Society of America, pp. 350-355, 1981.
15. CSL (Computerized Speech Lab):Kay Elemetrics Corp. 의

제품으로, 본 실험에서는 CSL Model 4300B를 이용하였다. 이것은 컴퓨터와 연결되는 하드웨어와 컴퓨터에서 실행 되는 음성분석 프로그램으로 구성되어 있으며, 이 프로그램은 아날로그(analog)나 디지탈(digital) 형태로 저장된 자료를 모두 이용할 수 있다. 각 음향요소의 측정시 화면에 표시된 파형에 눈금자(cursor)를 위치시키면 측정된 값이 화면에 수치로 나타나기 때문에 정확한 측정이 가능하다. 실제 분석에서는 컴퓨터의 화면을 음성파형과 피치를 나타내는 2 개의 창으로 나눈 후, 두 화면의 눈금자(cursor)를 하나로 연결시켜 피치값을 측정하였다.

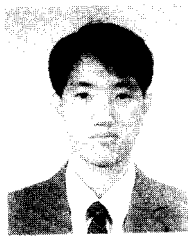
▲김 정 원 1969년 6월 14일생

1993년 2월 : 서울대학교 언어학과 학사
1995년 2월 : 서울대학교 언어학과 석사
1994년 9월~현재 : 디지콤 정보통신 연구소 연구원
※주관심분야 : 실험음성학, 음성합성

▲박 대 덕 1971년 3월 1일생

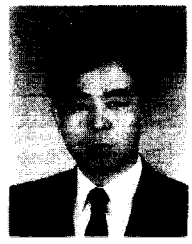
1993년 2월 : 홍익대학교 전자공학과 학사
1995년 2월 : 홍익대학교 전자공학과 석사
1995년 2월~현재 : 디지콤 정보통신 연구소 연구원
※주관심분야 : 음성부호화, 디지탈 신호처리, 음성합성

▲김 보 현 1969년 12월 28일생



1992년 2월 : 광운대학교 전자계산학과 학사
1994년 2월 : 광운대학교 전자계산학과 석사
1994년~현재 : 디지콤 정보통신연구소 연구원
※주관심분야 : 컴퓨터 네트웍, 인공지능, 음성합성

▲권 철 홍 1963년 7월 30일생



1987년 2월 : 서울대학교 전자공학과 학사
1989년 2월 : 한국과학기술원 전기 및 전자공학과 석사
1994년 8월 : 한국과학기술원 전기 및 전자공학과 박사
1989년 1월~1994년 8월 : 디지콤 정보통신연구소 연구원
1994년 9월~현재 : 디지콤 정보통신연구소 선임 연구원
※주관심분야 : 디지탈 신호처리, 음성정보 시스템, 음성합성