

## Signalyze에 의한 피치 분석방법 고찰

양 병 곤\*

### 목 차

- |                        |                        |
|------------------------|------------------------|
| I. 서론                  | V. 협역스펙트로그램을 이용한 피치측정법 |
| II. 성대에 의한 피치 생성       | VI. 결론                 |
| III. Signalyze의 피치측정방식 | 참고 문헌                  |
| IV. 정현파합성음의 분석값 비교     | Abstract               |

### I. 서론

사람은 얼굴표정이나 성대에서 발생되는 음원을 변화시켜 정서를 표현한다. 이러한 감정을 가장 잘 포착할 수 있는 방법은 성대의 떨림을 측정하는 피치 검출방법이 될 것이다. 그러나 컴퓨터에 의한 음성의 피치검출 방법은 다양하지만 각각의 방법에 대한 문제점과 결과치의 해석에 대해서 유의할 사항들이 있다. 사람의 발생원리를 무시하고 측정값에만 의존하거나 비싼 장비를 구입해 놓고도 사용법을 몰라 전혀 학술적으로 이용하지 못하는 경우도 많다. 또한 다양한 메뉴의 조합에 의해 결과치가 달라지는데 각각의 메뉴 설정에 대한 사전 지식이 없어서 연구 목적에 부합하는 결과를 얻는데 어려움이 있다. 따라서, 본 논문에서는 억양의 원리가 되는 성대의 작용에 대해 살펴보고 이를 Macintosh용 음성분석 소프트웨어인 Signalyze3.12의 다양한 메뉴 가운데 억양을 측정하는 도구인 피치측정방식을 사용하여, 합성된 정현파를 여러 가지 방식으로 분석한 결과를 비교해보고 가장 바람직한 분석방법을 찾아낸 뒤에 실제 화자의 음성에도 적용하면서 결과를 해석할 때 유의할 사항들을 고찰해 보고자 한다. 이 연구결과는 억양연구의 매우 기초적인 지식을 제공하여 다른 유사 음성분석프로그램에서도 쉽게 응용하여 기기활용의 극대화를 도모할 수 있으리라 여겨진다.

\* 동의대학교 인문대학 영어영문학과 부교수

## II. 성대에 의한 피치 생성

사람의 음성은 음원인 성대의 떨림을 여과기인 성도로 걸러서 원하는 목소리를 낸다고 한다(Fant 1960). 이 때 음원인 성대의 크기나 무게 질량 등이 중요한 변수로 작용하여 사람마다 독특한 음의 색깔이 정해진다. 일반적으로 성대의 작용은 공기역학적 근육탄성이론에 의해 설명한다. 즉, 폐에서 발생한 공기가 거의 맞붙은 성대의 공간을 밀고 열리게 하며 빠른 속도로 기류가 흘러가면서 기압이 낮아지며 일종의 진공상태가 되어 양측의 성대가 서로 맞붙게되고 이런 여닫음이 1초에 성인 남성의 경우에는 120번 정도로 진동하고, 여성인 경우 230번 어린애일 경우는 180-500번이나 자주 떨리게 된다(Boothroyd 1986). 성대의 떨림을 나타내는 기본 주파수는 여러가지 요소에 영향을 받는다. Eguchi와 Hirsh(1969)는 3세에서 6세 사이에 기본 주파수가 급격히 감소하다가 그 후 서서히 변해서 13세 경에는 여성은 성인의 주파수에 도달하고 남자는 계속 감소해 나감을 밝혔다. 성대길이는 탄생시 3mm에서 두 살에 5.5mm 여섯살에 8mm 15세에 9.5mm이고 성인 여자의 성대길이는 12.5에서 17mm이고, 성인 여자의 경우는 17내지 23mm이다(Negus 1949). 기본 주파수는 성대의 무게와 길이에 반비례하고 긴장에 비례하므로, 보다 작은 성대를 가진 여성의 성대진동은 남성보다 빠르다고 볼 수 있다. 기본 주파수는 또한 화자가 후두 근육 긴장도를 조절하거나 성문하기압을 조절하여 바꿀 수 있다. 해부학적으로는 성대는 갑상연골과 피열연골을 조절하는 근육을 수축하여 피열연골이 기울어져서 성대를 당기게 함으로써 긴장을 증가시킬 수 있다. 이러한 성대의 떨림 정도를 청각적으로 지각하는 잣대를 보통 피치라 부른다. 연속적인 피치의 변화를 억양변화곡선으로 나타내어 발화자의 감정변화나 숨겨져 있는 의미를 알 수 있기 때문에 일상대화나 음성언어학의 연구분야에서는 중요한 위치를 차지하고 있다.

성대의 작용에 대해서는 음성병리학자들에 의해 많은 연구를 하고 있으며 악성종양 등의 진단과 치료후 재활 정도에 대한 측정값을 구하기 위해 PGG(Photoglottography)를 통해 성대사이의 빛의 투과성 정도를 측정하거나, EGG(Electroglottography)를 이용해 성대사이의 전류의 흐르는 양 또는 저항정도를 측정하는 방안이 제시되어 있다(양병곤 1996:1025-40 참고). 그러나 이러한 장비는 마취를 하여 전극이 달린 주사바늘을 피험자의 목에 정확히 주입하여 성대의 떨림을 측정할 수 있다. 특히, 우리의 일상생활에서 흔히 사용되는 억양변화를 측정하기에는 너무 비싼 비용을 지불해야한다. 사전 실험에 의하면 언어학적인 억양변화의 측정은 오히려 간단한 Signalyze의 피치 측정방식을 이용할 수 있다. 아울러 좁은 대역의 스펙트로그램을 이용하면 더 정확히 측정할 수 있다.

## III. Signalyze의 피치측정방식

Signalyze에는 그림 1과 같은 피치측정 설정에 관한 대화창이 있다.

여기에는 신호에서 성대의 떨림 속도를 측정하기 위한 세 가지 측정방식이 나온다(Keller 1992). 첫번

째는 시간 구조 분석(Temporal Structure Analysis:TSA) 방식으로 신호가운데 성대의 떨림을 나타내는 것은 가장 큰 진폭으로 규칙적인 간격으로 떨어져 있다는 점을 이용한다. 이 방식은 녹음이 잘된 음성에서 사용하지만 한 옥타브이상의 큰 피치 변화(예를 들어, 90Hz를 기준으로 하면 그 두 배인 180Hz이상의 범위)에는 사용하지 않는다. 두번째는 FFT-comb 방식으로 일정 범위의 신호창을 고속푸리에변환(Fast Fourier Transform:FFT)을 하여 스펙트럼상 규칙적으로 나타나는 배음을 포착하여 피치를 구하는 방식으

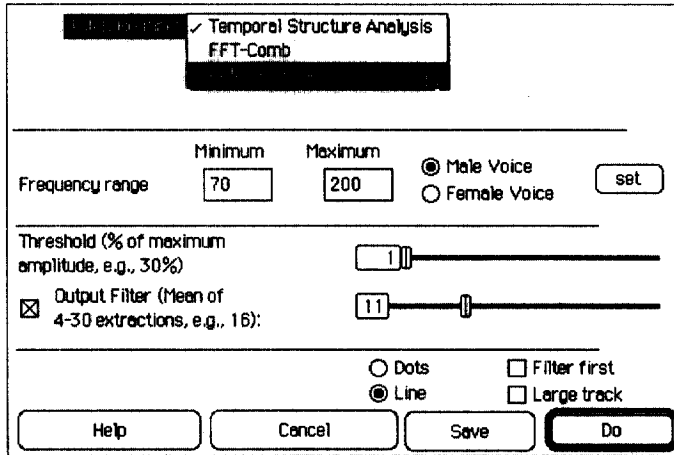


그림 1. 피치 측정 설정방식 대화창

로 여러가지 유형의 음성신호에 사용할 수 있으나 90Hz이하의 낮은 음성에서는 좋지 않다. 세 번째는 자기상관법(Autocorrelation:AC)으로 성대의 떨림을 나타내는 신호의 한 사이클이 따라오는 유사한 사이클과 높은 상관도를 보인다는 점을 이용한다. 이것은 낮은 주파수나 소음이 많은 경우에 사용하나 느리고 주파수 범위를 세밀하게 정해줘야 하는 단점이 있다. 불규칙적이고 빠른 성대 변화는 여과시켜 평균값을 주기 때문에 비정상적인 음성에는 부적합하다. 여기서, FFT-comb이나 AC 방식에서는 스펙트럼 측정 범위를 지정해야하는데 보통 5ms에서 10ms로 지정한다. 범위가 좁을수록, 고주파의 음을 정밀하게 측정하는데 사용할 수 있다. 각각의 방법에서 주파수 범위(Frequency Range)를 지정할 때 주의해야 한다. TSA에서는 한 옥타브이내의 값을 지정해줘야하고 FFT-comb에서는 최소값이 계산에 이용되고 최대값은 창에 나타낼 때만 이용된다. AC에서는 최대값과 최소값 모두가 계산에 이용되나 TSA보다는 값을 자유롭게 지정할 수 있다. 결과 창은 이 범위내에 해당하는 값만을 걸러서 나타내 주기 때문에 편리하다. 남여목청 지정은 미리 정해진 최대값과 최소값을 기본값으로 빠르게 전환할 때 사용된다. 이것을 먼저 선택한 뒤 적정범위의 값으로 변경한다. 경계값(Threshold)은 컴퓨터가 피치분석을 할 때, 목청떨림이 있고 없음을 판단하는 값으로 이 값을 기준으로 분석을 할 것인지 그냥 지나갈 것인지를 결정한다. TSA에서는 신호의 기울기를 기준으로하고, FFT-comb은 강도와 저주파기준을 사용하고 AC는 강도를 기준으로 삼아 보통 음성분석구간에 속해있는 파형의 진폭값이나 기울기의 최대값을 구한 뒤 그 값의 10-20%범위의 값을 유무성구분의 범위로 삼아서 분석을 한다. 이 값을 너무 높이면 일부 분석 결과만 보이게 되고, 너무 낮게하면 유성자음의 떨림까지도 분석하여 오류치가 나오게 된다. 밑에 따라오는 출

력여과기(Output Filter)는 명백히 벗어난 피치값을 없애주는 역할을 한다. 구체적으로는 지정된 갯수 만큼의 표본값 가운데 지정한 주파수값 아래위 25%범위의 값을 버리고 남아있는 50%의 표본값 중 평균 주파수를 계산해 준다. 따라서, 계산 방식상 4의 배수값(4, 8, 12, 16 등)을 지정하는 것이 좋으며, FFT-comb이나 AC 방식에서 5 ms마다 계산될 때는 8개의 표본을 선택한다. TSA에서는 12, 16, 20 등의 값을 써도 되지만 FFT-comb이나 AC 방식에서는 넓은 분석구간(10 ms 이상)과 8보다 큰 여과기값을 지정하면 시간해상도가 떨어진다. 분석된 결과는 점과 선으로 나타낼 수 있고 고주파부분의 간섭을 피하려면 미리 음성신호를 저주파통과 여과기로 처리하는게 좋다. 분석구간은 남자일 경우에는 5-7ms를 사용하고 여자일 경우는 그보다 좁은 구간을 지정한다. 구간이 좁을수록 계산시간이 더 걸린다.

#### IV. 정현과합성음의 분석값 비교

그림 2는 SoundEdit16의 메뉴에 있는 정현과합성기로 합성한 음파를 분석한 것이다. 맨 위의 창은 사람의 피치값에 가까운 100Hz, 150Hz, 200Hz, 250Hz의 네 개의 음을 각 200ms간격으로 만들었고 이들 사이의 경계를 20ms로 두고 표본속도는 22250Hz로 합성했다. 이 합성음을 피치분석한 것은 이어지는 창마다 다른 방법을 지정했는데, 먼저 TSA와 5ms창으로 FFT-comb, Autocorrelation방식을 지정했다.

창문의 x축 주파수 범위는 80-300Hz범위로 했다. 먼저 TSA방식에서는 200Hz부분이 150Hz로 약간 포착이 되었다가 바로 66Hz로 내려갔다. FFT-comb 방식에서는 250Hz만이 거의 비슷하게 측정되었으나 나머지는 변화가 많았다. Autocorrelation방식에서는 거의 적절히 측정이 되었으나 저주파인 100Hz는 전

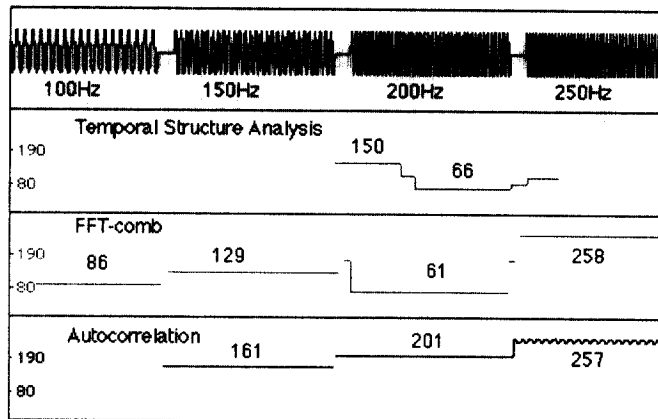


그림 2. 정현과합성음에 대한 피치분석 결과

혀 측정값이 나오지 않았다. 각각의 처리방식에 문제가 있음을 극명하게 보여주는 예가 될 것이다. 따라서, 이 프로그램을 이용할 때는 측정된 파형에 대해 보다 비판적으로 접근해야 할 것이다.

보통 피치값의 정확성을 눈으로 확인하기 위한 가장 좋은 방법은 그림 3과 같이 신호를 확대하여 제

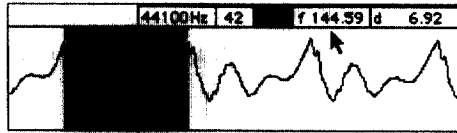


그림 3. 피치주파수를 확인하는 방법

일 큰 강도의 위치사이를 선택한 뒤 주파수(f)를 나타내는 창문의 값을 읽으면 된다. 여기서는 144.59Hz가 된다. 이 값은 1000밀리초안에 바로 옆 칸의 지속시간(d) 6.92 밀리초의 간격으로 몇 개의 사이클이 존재하는가를 구하는 것과 마찬가지로  $1000/6.92=144.59$ 로 구해진 것이다. 마우스로 정확히 구간을 잡지 못했기 때문에 생긴 오차가 있으나, 실제 분석값의 10자릿수 이상의 값이 동일하게 포착되었으면 정확하다고 판단해도 된다.

다음으로는 40세 남성과 11세의 여자아이가 “이것은 책입니다”라는 문장의 단어 “책”을 임의로 강조점을 두어 발음한 음성을 분석해 보고자 한다. 이들의 피치는 아주 낮고 높은 대조를 이루기 때문에 음성분석기의 성능을 시험하는데 도움이 되기 때문에 선정했다. 음성입력은 11125Hz의 표본속도로 마이크로 PowerPC에 입력하여 분석해 보았다. 그림 4의 맨 위의 창문은 음성파형을 나타내고 있으며 두 번째 파형은 TSA분석을 이용했고 y축의 주파수 범위는 100-300Hz로 했다. 두 번째 곡선은 FFT-comb 방식에 의한 분석으로 분석구간은 5ms마다 실행되도록 했고 문턱값(Threshold) 설정은 10%로 지정했다. 세 번째 창문은 이 음성을 5ms마다 AC 분석을 이용했다.

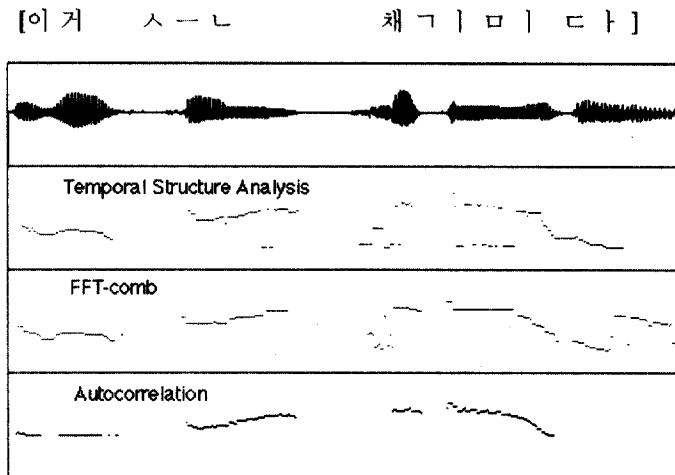


그림 4. 40세 남성화자의 발성음에 대한 억양 분석곡선

이 그림에서 살펴보면 두 옥타브를 넘는 범위를 지정했기 때문에 TSA는 매우 불안한 억양패턴을 보이고 있다. 두 개의 억양패턴이 동시에 형성되는 듯한 이중피치모양을 보이고 있다. 반면 FFT-comb 방식에서는 [은]과 [책] 부분에서 변화를 포착하지 못하고 수평으로 나타나있으며, 마지막 [다]에서는 실제로 떨어져야 하는데 오히려 피치가 올라간 모양을 보이고 있다. 마지막 창문의 Autocorrelation에서는 맨처

음의 [이]와 따라오는 [거시] 이 수평으로 나타나 있다. 마지막 음절인 [다] 부분은 아예 포착이 되지 않았다. 이러한 자료를 평균화시킨다면 오류가 더 심해질 것임이 명백하다. 세 가지를 비교한다면 그래도 FFT-comb에서 모양이 잘 포착이 됐다고 볼 수 있다.

다음으로는 11세 여아의 음성을 동일한 조건으로 입력하고 분석해 보았다. 다만 피치범위는 남성화자의 범위보다 넓기 때문에 150-400Hz로 설정했다. 그림 5에서 살펴보면 TSA는 두 옥타브 범위를 넘어서기 때문에 음절 [은]과 [책]에서는 두개의 피치패턴을 보여주고 있다. FFT-comb 방식에서는 5ms간격 인데도 불구하고 동일한 높이로 측정된 모양이 많이 나타난다. 마지막 음절 [다]는 문장의 끝으로 피치가 내려가는 부분인데 올라가는 모양으로 끝나고 있다. 마지막 음절인 [다]에서는 TSA와 동일한 패턴을 보여주고 있다. 자기상관법에 의한 파형은 자음 [기], [치], [디]에서 잘못된 피치패턴이 추출되어 상한값인 400Hz에 분포되어 있고 끝 음절 [다]는 아예 분석값이 나타나지 않는다.

이러한 결과는 앞서의 각 분석방법의 문제점을 잘 나타내주고 있으며 남성화자와 여자아이의 억양패턴을 그래도 적절히 포착한 것은 FFT-comb이 될 것이다. 실제 FFT를 한 뒤 생긴 스펙트럼을 음성파형으로 간주하여 한번 더 FFT를 실행하는 Cepstrum을 사용하면 거의 정확히 피치 배음을 측정할 수 있을

[이 거 시 ㄴ 책 기 | ㅁ | 디 다]

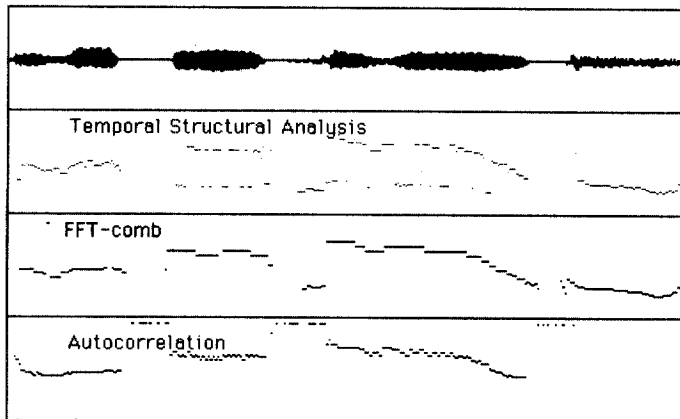


그림 5. 11세 여아의 발음에 대한 억양 분석곡선

것이다. 협역스펙트로그램을 이용하여 피치변화를 측정할 예로는 Koo(1986:38이하)가 모음 부분에 해당하는 협역스펙트로그램 분석에서 나타난 열 번째 배음의 값을 추정할 뒤 이것을 10으로 나누어 값을 구하거나 열 번째 배음이 나타나지 않을 때는 다섯 번째 배음을 이용해서 구했다. 또한 Kent(1992:79 그림4-12 참조)는 협역스펙트로그램의 첫 번째 배음이 보이는 0-250Hz 부분만 확대하여 억양 변화를 표시하고 있다. 따라서, 본 논문에서는 FFT를 통해 배음구조를 상세히 보여주는 협역스펙트로그램을 이용하여 억양구조를 나타내는 방법을 찾아보고자 한다.

### V. 협역스펙트로그램을 이용한 피치추정법

Signalyze에서는 열 가지 종류의 스펙트럼을 분석할 수 있다. 이들 가운데 피치 분석에 활용할 수 있는 방식을 살펴 보면 먼저 아주 가는(Very Slim) 대역은 10Hz 폭으로 100ms분량의 신호를 훑어나가면서 공명정도를 측정하여 화면에 표시해 주며, 음악이나 동물의 소리를 분석할 때 사용한다. 가는 대역(Slim)은 50ms 분량의 신호를 20Hz 간격으로 훑어가는데, 음악이나 동물의 소리에 사용할 수 있고 남자의 목소리 분석에 편리하다. 아주 좁은(Very Narrow) 대역은 22ms의 신호를 20Hz 폭으로 분석하는데 남자의 피치분석이나 배음주파수를 구하는데 적합하다. 좁은(Narrow) 대역은 25ms의 신호를 40Hz 폭으로 분석하며, 여성이나 아이들의 피치와 배음주파수를 구하는 데 편리하다.

앞서의 합성한 정현파를 두 가지 방식에 의해 분석해보면 그림 6과 같다. 그림 6의 두 번째의 창에 생성된 스펙트로그램은 33ms에 해당하는 구간에 대해 30Hz간격의 아주 좁은 대역으로 분석하고 이를 나타내는 범위로 표본주파수의 1/4 유효영역에 해당하는 1378Hz까지 나타내었다. 피치의 변화가 화면 아래 깔려있어서 그 크기를 상대적으로 구분하기에는 불편하다. 보통 피치값의 측정에서도 고주파부분이 분석에 장애를 주기 때문에 저주파영역만 측정범위로 하기위해 저주파통과 여과기를 사용한다. 이것을 선택하면 정상적인 사람의 음성인 피치범위내에 해당하는 주파수, 예를 들어 500Hz이하의 음만을 통과시킨다. 따라서, 이번에는 22250Hz의 표본 속도를 1000Hz로 변환하는 복합변환(Complex Transformation)을 실행한 뒤 이를 다시 한번 아주 좁은 대역으로 처리하여 유효영역인 500Hz까지 나타내 보았다. 이것은 매우 선명하고 상대적으로 깨끗한 피치패턴을 보여준다.

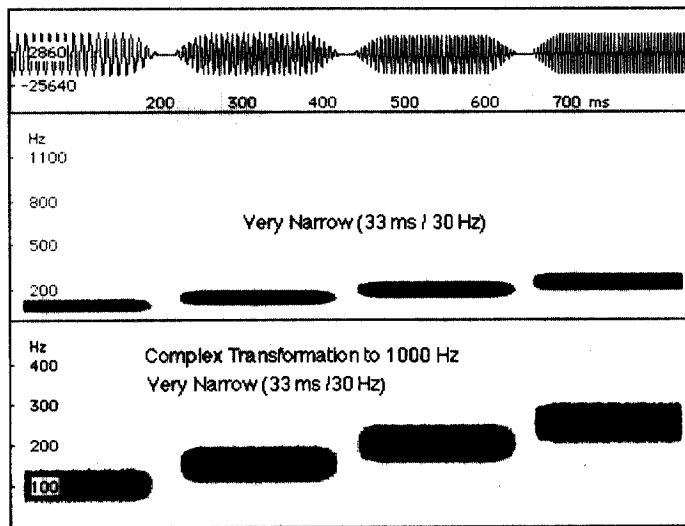


그림 6. 협역 스펙트로그램을 이용한 피치 분석법

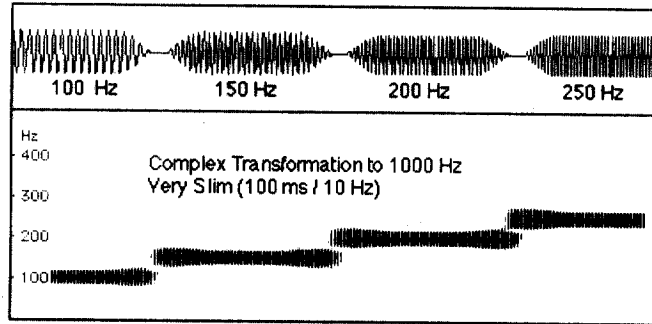


그림 7. 정현파 복합음을 복합변환을 통해 변환한 뒤 아주 좁은 대역으로 분석한 결과

특히, 앞서의 여러가지 방식에서 주파수측정값이 지정범위에 따라 분석의 문제점으로 지적되었던 정확한 수치값이 여기서는 모두 바르게 상승하는 피치의 변화를 포착하고 있음을 관찰할 수 있다.

다음으로는 복합변환을 한 다음에 보다 더 얇은 주파수대역(100 ms/10 Hz)을 사용하고 스펙트럼 표시는 FFT분석한 결과를 하나씩 걸러 나타내는 방식(Plot every other line)을 이용해 보았다. 그림 7은 앞서의 분석에서 나온 진한 그림 6보다는 열게 되어 여기에 중심선을 임의로 그어 보았다. 이 배음의 두께가 더욱 더 가늘어져서 배음간 간격을 넓혀주어 더 선명한 억양변화를 포착할 수 있다. 그러나 분석구간이 100ms에 걸쳐있기 때문에 신호간 사이에도 앞 뒤의 신호값에 의한 흔적들이 이어져 있음을 볼 수 있고 마지막 신호의 분석구간은 실제 신호보다 이전에 배음곡선이 끝나도록 되어 있으므로 경계선 사이의 분석결과에 대해서는 해석할 때 주의해야 한다. 특히, 자음 구간이 아주 짧은 경우에는 분석구간

[이 거 사 - 나 채 가 | 모 | 다 | ]

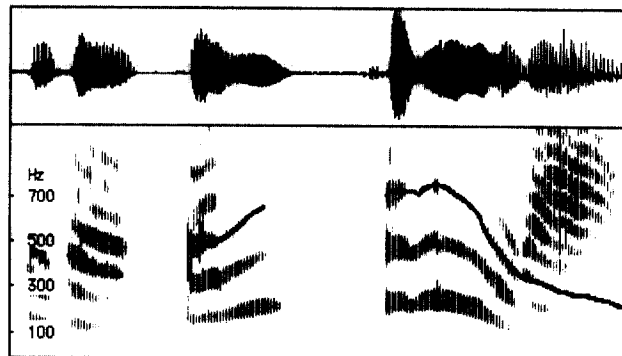


그림 8. 40세 남성의 억양변화

이 넓어서 실제 목청떨림이 없는 것임에도 불구하고 연속된 배음으로 분석될 우려가 있다.

그림 8은 앞서 녹음한 40세 남성의 목소리와 11세 여아의 목소리를 복합변환을 시켜 3000Hz로 표본



속도를 내린 뒤 이를 33ms간격마다 30Hz의 좁은 대역으로 분석한 스펙트로그램이다.

이 그림에서 살펴보면 남성화자의 억양곡선이 1500Hz까지 나타내는 척도로 되어있고 [은] 부분부터 점점 올라가서 [채깅]에는 최고조로 도달했다가 하강한다. 편의상 세 번째 배음위에 중심선을 그어 억양의 변화를 눈으로 쉽게 확인할 수 있게 했다. 첫 번째 배음의 변화를 보면 약 120Hz 전후의 변화를 보여주고 있다. 최고조로 상승한 지점의 주파수는 약 230Hz까지 올라가 있다. 특히, [채기미다] 부분은 계속적인 배음연결로 유성화된 모습을 보여주고 있다.

그림 9는 11세 여아의 발성음의 억양변화를 분석한 그림이다.

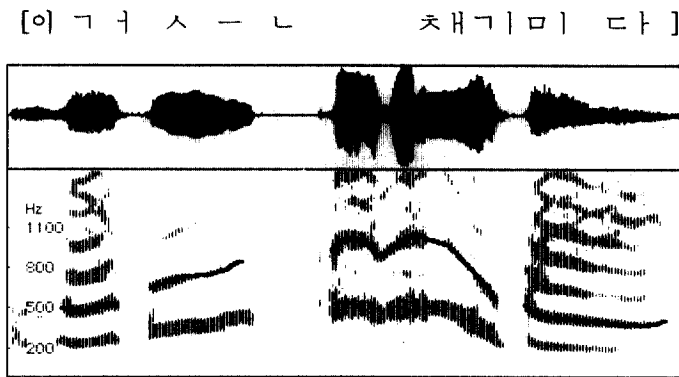


그림 9. 11세 여자아이의 억양변화

그림 9에서 살펴보면 제1배음이 200Hz를 넘어서는 높은 피치의 음성임을 쉽게 알 수 있다. 아울러 강하게 억양이 주어질 때는 약 430Hz나 높이 올라갔다가 다시 내려오는 모습을 보여준다. [채기미] 구간에서는 두 개의 산봉우리를 형성하고 있으며 마지막 음절 [다]는 무성음으로 연음이 되지 않음을 나타내 준다. 그림 8의 남성의 스펙트로그램보다 배음의 간격이 매우 성기게 나타나 기본 피치가 높음을 알 수 있다. 협역스펙트로그램에서는 앞서 살펴본 여러가지 피치분석에서 발생하는 소음이나 기타 억양을 포착하는데 장애가 별로 없어 보인다.

## VI. 결론

지금까지 Signalyze를 이용하여 사람의 정서를 표현하는 성대의 떨림속도를 측정하는 방법에 대해 살펴보았다. 이 소프트웨어는 소형컴퓨터에서 실행할 수 있는 간편하면서도 강력한 분석메뉴를 갖추고 있다. 그러나 지금까지 단순한 정현파의 분석에서 살펴보았듯이 여러가지 문제점을 갖고 있으며, 이러한 문제점을 극복하기 위해서는 오히려 좁은 대역의 스펙트럼분석을 활용한 스펙트로그램의 궤적을 이용하는 것이 정확한 피치변화를 추적하는데 도움이 됨을 알 수 있었다. 아울러, 높은 표본속도를 화자의 피치변화폭의 범위의 두배에 해당하는 낮은 표본속도로 복합변경을 통해 내린 다음에 이것을 협역스펙

트로그램으로 분석함으로써 저주파부분의 배음이 더 확대되어 나타나고, 그 중심선을 그어나가면 보다 나은 피치곡선을 얻을 수가 있었다. 앞으로 이러한 기법을 응용하여 국어억양 및 영어억양에 대해 보다 많은 연구가 활성화되기를 기대해본다. 앞으로 억양 측정에 관한 타기종의 컴퓨터 프로그램에 대해서도 이와 비슷한 검증과정을 거쳐 새로운 분석방법을 찾아볼 계획이다.

## 참고 문헌

- 양병곤. 1996. 라리고그래프에 의한 한국인의 성문과형 분석 및 합성모음의 청각실험. 언어 21권 4호 pp. 1025-1040.
- Boothroyd, A. 1986. *Speech Acoustics and Perception*. Austin: Pro-Ed, Inc.
- Eguchi and Hirsh. 1969. Development of speech sounds in children. *Acta Oto-Laryngologica*, 16, pp.1-51.
- Keller, E. 1992. *Signal Analysis for Speech and Sound: User's Manual*. Charlestown, M.A.:Infosignal Inc.
- Kent, R.D and C. Read. 1992. *The Acoustic Analysis of Speech*. San Diego, CA: Singular Pub. Inc.
- Negus, V.E. 1949. *Comparative Anatomy and Physiology of the Larynx*. New York: Grune & Stratton.
- Ko, D.H. 1988. *Declarative Intonation in Korean. An acoustical study of F0 declination.*. Seoul: Hanshin Pub. Inc.
- Koo, H.S. 1986. *An Experimental Acoustic Study of the Phonetics of Intonation in Standard Korean.*. Seoul: Hanshin Pub. Inc.
- Yang, B. 1990. *Development of Vowel Normalization Procedures: English and Korean*. Seoul: Hanshin Pub. Inc.

<Abstract>

## A Study of Pitch Analysis Methods by Signalize

*Yang, Byung-gon*

The purpose of this study is to propose a better method of getting pitch pattern of speech sounds. Pitch is a perceptual scale from which people catch emotional variation of a speaker involved. Therefore, it plays an important role in everyday conversation as well as in speech analysis. It is normally produced by the changes in fundamental frequency or vocal fold movement. First, the pitch analysis methods of Signalize are critically reviewed. Each method is tested on a complex tone of 100-250Hz mixed by a sound software to check its problems and limitations. Also, pitch variation is observed in two subjects of different pitch. They produced a Korean sentence emphasizing a word in a sentence with higher pitch. Pitch contour is traced on a spectrogram produced by setting a very thin band after downsampling the signal data at the rate of the subjects' pitch range doubled. Results show very clear tonal variation traced along each mid point of the third harmonic, which supersedes the three pitch analysis methods of Signalize.