

## 16. Pitch, Periodicity, and Noise in the Voice

Perry R. Cook

### 16.1 Isn't Singing Just Strange Speech?

#### 말하기와 비교한 노래 발성의 특징

항 목	특 징
성음/무성음 비율	성음비율이 지속적으로 95% 유지
비브라토	의도적으로 발생하는 음고의 편차
포먼트	공명의 증폭을 위해 4~5개 이상의 포먼트가 함께 나타나는 현상
모음의 변형	의도적, 또는 비의도적으로 모음의 발성을 변형
Nasal Airway의 사용	말할 때만큼 많이 사용되지 않음
음역대와 평균 음고	말할 때 편한 음고와 노래할 때 편한 음고는 서로 다름 가수의 음역대는 생리학적 요인과 훈련에 의해 결정됨
음량 범위와 평균 음량	말할 때의 평균 음량은 노래할 때의 평균음량에 비해 작음 말할 때의 다이내믹 레인지는 노래할 때에 비해 낮음
발성 훈련	성대를 훈련함으로써 가수가 아닌 사람들과 다른 음원을 갖게 됨
신경학	노래와 말하기를 관장하는 뇌의 영역이 서로 다른 것으로 추측됨
법적인 관점	말하기의 자유는 헌법에 보장되어 있으나 노래하기의 자유는 보호되지 않는 경향이 있음

### 16.2 Pitch in Singing

#### 사람의 목소리와 일반 악기들 간의 비교

종류	참조 음고(Reference Pitch)	연속된 음고(Sweepable Pitch)
사람의 목소리	없음 : 변하는 상황에 맞추어 근육을 조절하여 발성	가능
악기 (Continuum of Pitch)	있음	
악기 (Discrete Pitch)	: 악기의 물리적 형태에 의해 참조 음고가 결정됨	불가능

- 절대 음감을 지닌 사람들이라 할지라도 훈련된 가수가 아닌 경우에는 생각하는 음을 정확하게 내는 것이 불가능하다.
- 절대 음감을 지닌 가수 역시 나이, 질병 등의 요소에 따라 발성에 영향을 받으며, 특히 자신이 내는 소리를 들을 수 없는 상황에서는 정확한 음고의 발성은 불가능하다.

(Ward and Burns 1978)

**비브라토**

- 예술적인 목적으로 발생시키는 의도적인 음고의 변조
- 일반적으로 5~8Hz의 범위의 주파수 편차를 가짐
- 비브라토를 의도하지 않더라도 측정 가능한 작은 양의 비브라토가 존재 [그림1]

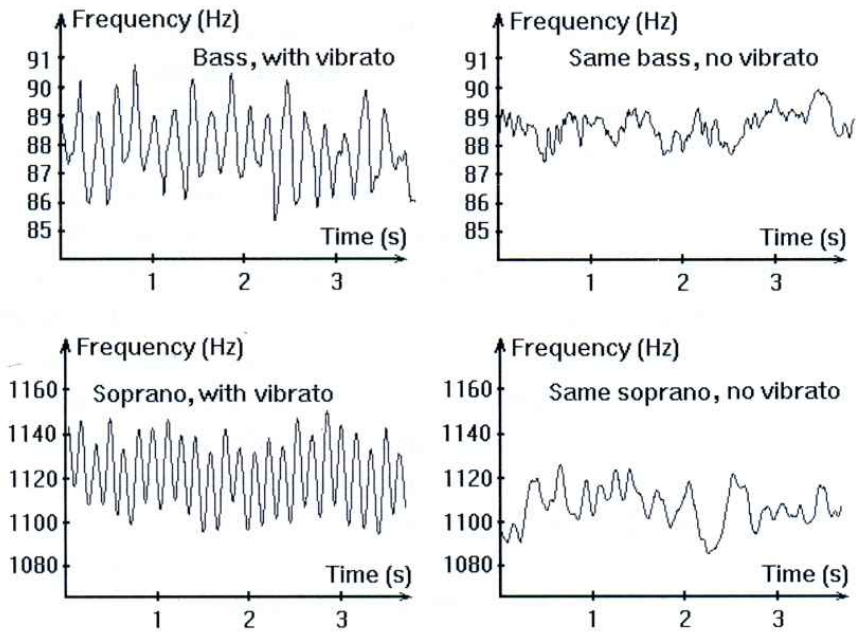


그림 1 비브라토를 의도한 발성 및 비브라토를 의도하지 않은 발성의 분석

- 비브라토 발성 시의 음고 만을 푸리에 변환을 통해 분석한 경우, 6Hz의 고점이 뚜렷하게 발생하며, 비브라토가 없는 발성에서도 6Hz의 고점은 발생 [그림2]

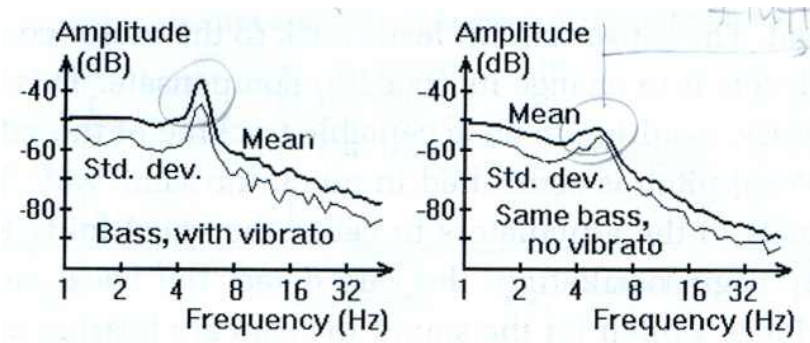


그림 2 비브라토의 유무에 따른 6Hz 고점 발생에 대한 분석

- 6Hz 고점을 중심으로 윗부분을 flutter, 아랫부분을 wow 영역이라 부름  
 : 레코딩 시에는 이 영역들의 음량이 작을수록 좋으나, 인간의 목소리를 컴퓨터로 합성할 때에는 이 요소들의 재현이 중요

**Jitter (Flutter)**

- 의도적 제어가 불가능한 생리학적 떨림 현상
  - : 물컵을 권 손이 떨릴 때, 이 떨림을 막으려 더 힘을 주면 오히려 더 떨리는 현상
  - : 조음을 제어하는 근육에서도 이와 비슷한 빠른 떨림이 발생(고음역대에서 주로 발생)

**Drift (Wow)**

- 피드백 기제에 의해 진행되는 오차수정과정에서 발생하는 오류
  - : 인공위성과 지구의 송신기 사이의 위치 추적과정에서 찾아볼 수 있음
  - : 발성 - 귀로 듣고 음고 및 음량 판단 - 발성형태 수정 - 다시 귀로 듣고 판단...
- Drift는 Jitter에 비해 제어가 쉽게 가능함 - 청각 피드백의 역할이 중요
  - : 피드백의 상태가 좋지 않은 경우 Drift의 양은 증가함
  - : 저음 발성에 있어 Drift가 증가하는 경향이 있음(저음의 음고 감지가 더 어려움)
- Jitter는 청각 피드백에 크게 의존하지 않으나, 비브라토가 존재하는 경우 증가함
- Drift는 목소리의 음량과는 거의 무관하며, Jitter는 음량이 증가할 때 증가하는 경향을 보임 [그림 3]

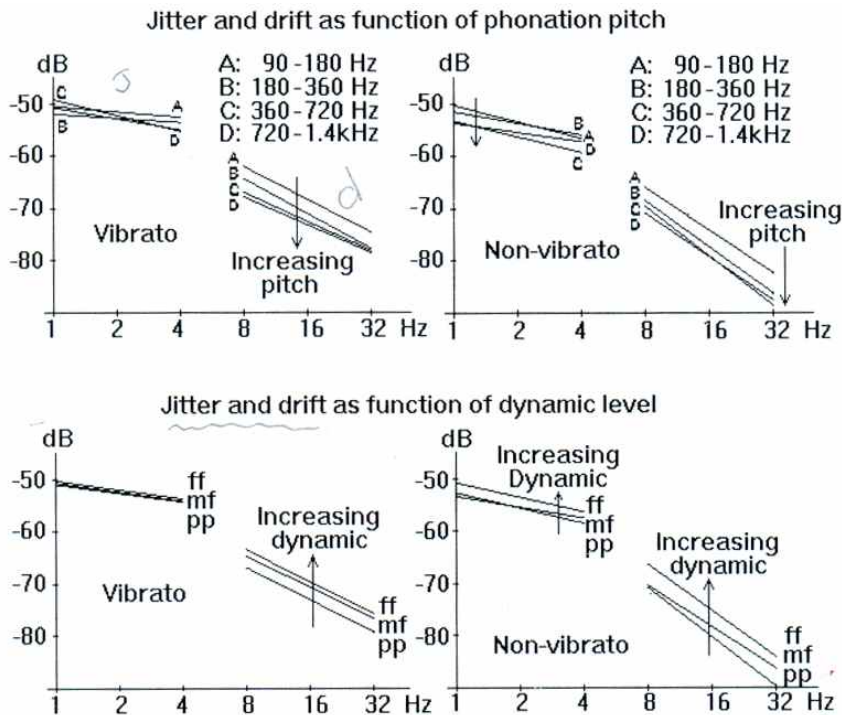


그림 3 Jitter, Drift - Amplitude, Pitch, Vibrato간의 관계

- 일반인들은 자연되거나 변화된 청각 피드백을 듣게 될 경우 비브라토 발성에 영향을 받음 (Deutsch and Clarkson 1959)
- 훈련된 가수들은 왜곡된 청각 피드백에도 거의 영향을 받지 않고 비브라토 발성을 할 수 있으나, Drift의 양은 증가하는 경향을 보임 (Shipp et al. 1984)
- 서양음악에서의 비브라토 정도의 범위는 ±2%로, ±4%를 지나치면 어색하게 들림
- Drift의 범위는 0.5% 정도이며, Jitter의 범위는 8Hz일 때 0.05%, 32Hz일 때 0.005%까

지 감소함 [Trk. 71]

- : Drift가 큰 경우 가수가 자신의 목소리를 듣지 못하는 것과 같은 효과를 내며, Jitter가 큰 경우 가수가 정신병을 앓고 있거나, 조악한 녹음기로 녹음된 것과 같은 효과가 발생
- 음성의 비주기성(aperiodicity)에서 나타나는 현상을 Shimmer라고 하며, 이로 인해 음색의 변화가 일어남

16.3 Noise in the Voice

- 성대(vocal folds)의 열고 닫히는 물리적인 움직임과 그 사이를 통해 배출되는 공기에 의해 잡음이 발생할 가능성이 충분히 있음
- 노래의 발성을 통해 나오는 소리의 파형은 기본적으로 주기성을 갖고 있으나, 한 주기 주기를 미세하게 들여다보면 서로 다름 [그림 4, 5]

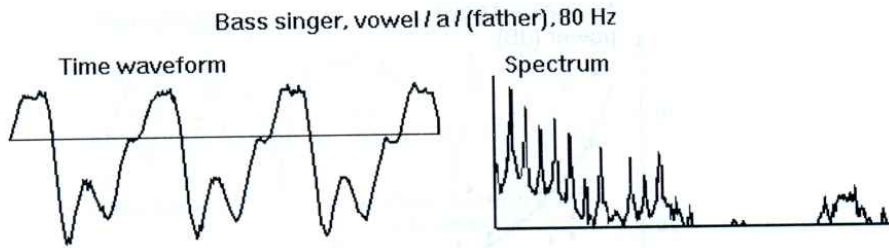


그림 4 남성의 /a/ 발성의 Time Domain 분석과 스펙트럼 분석

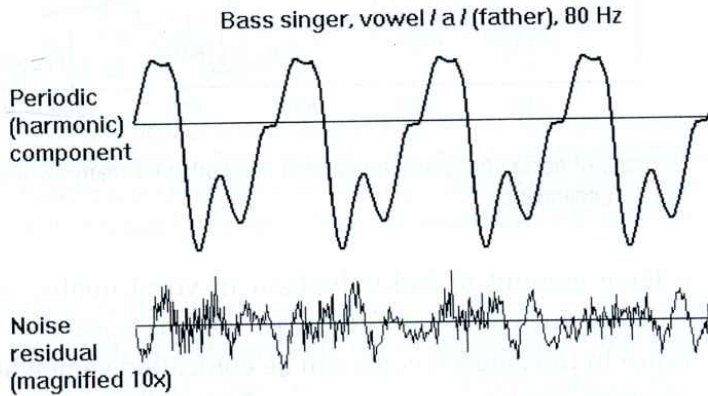


그림 5 신호의 분해(decomposition)후 Harmonic 성분과 Noise Residual의 비교

Voice Quality

일반적으로 음성으로 얻을 수 있는 중요한 정보는 음고, 음량, 발성하는 단어 등이 있으나 음성의 주파수 분포 경향, 잡음의 양, 포먼트의 위치에 의해 Voice Quality가 정해짐

- 발성시 음고가 높아질수록 잡음의 양은 감소하는 경향을 보임(반비례 관계)

16.4 Subharmonics in the Voice

- 발성을 의도한 음의 음고외에도 다른 음고를 가진 소리가 포함되어 있으며, 특히 1/2,

3/2배의 주파수 관계를 가진 배음들이 나타나는데 이를 Subharmonics라 한다.  
 [그림 6]

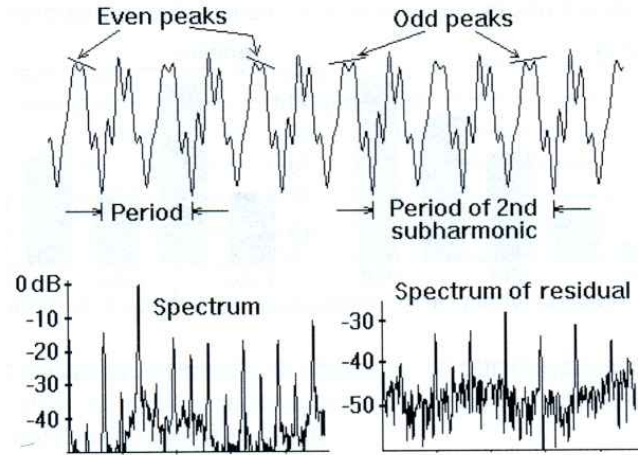


그림 6 340Hz의 남성의 음성에서 발견되는 Subharmonics

**Diplophonia**

소리의 파형이 가진 원래의 주기와는 다른 형태로 또 다른 주기(pseudo-period)가 발견될 때, Subharmonics가 발생하며, 이러한 현상을 Diplophonia라 부르기도 한다.

- 20명의 가수를 대상으로 한 Subharmonics 연구 (Cook 1991)
  - 20명 모두 측정가능한 정도의 Subharmonics가 발견되었음
  - Subharmonics를 포함하고 있는 소리의 비율: 남자 34%, 여자 32%
  - 모음의 발성에 의한 비율: /a/ 34%, /u/ 40%, /i/ 20%
  - 음량에 의한 비율: 약함 10%, 중간 38%, 강함 20%
  - 남성의 경우 명확한 패턴을 가지고 있으나, 여성의 경우 패턴이 없음 [그림 7]

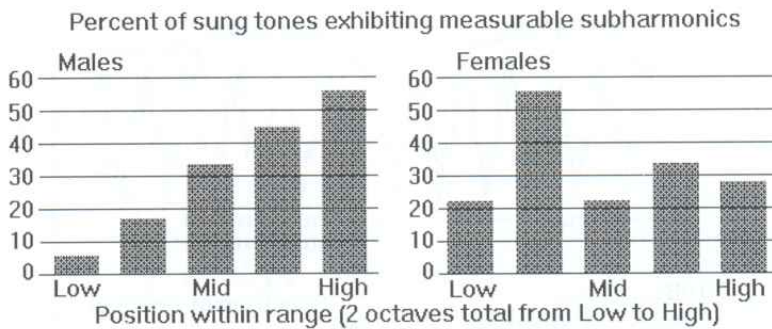


그림 7 성별에 따른 분석: 음역대와 측정 가능한 Subharmonics의 상관관계

**16.5 Pitch and Noise in Vowel Perception**

- 높은 음고의 발성에서는 모음의 발음이 대부분 비슷하게 들림  
 : "All sound the same"
- 음고가 높아짐에 따라 음색을 구성하는 배음 간의 간격이 벌어지며, 동시에 포먼트의 간격도 벌어지게 되어 모음의 특색을 나타내는 포먼트의 형태를 듣고 판단하기 힘들
- 훌륭한 작사가로 유명했던 작곡가들은 이러한 점을 이해하고 있었으며, 중요한 가사를

낮은 멜로디에 배치했다. 가사가 높은 음에 배치되어야 하는 경우에는 앞서서 낮은 음에 가사를 미리 등장시킴으로서 청중들에게 가사를 각인시키는 기법을 사용 (Scotto di Carlo and Rutherford 1990)

### Hidden Formant

- 넓은 비브라토 등에서 나타나는 음고의 편차로 인해, 청자는 음색 형태를 보다 잘 이해하게 된다고 주장되어 왔음

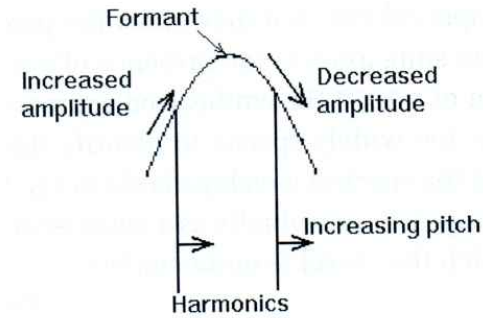


그림 8 Hidden Formant

: 왼쪽의 배음은 주파수가 상승할수록 음량 증가, 오른쪽의 배음은 주파수가 상승할수록 음량이 감소; 결과적으로 두 배음 사이에 포먼트가 존재함을 지각 - 비브라토에 의해 포먼트가 더 잘 들리게 되고, 이에 따라 모음의 지각이 더 쉬워진다는 주장 [그림 8]

- Johann Sundberg(1978)은 실험을 통해 위의 주장이 사실이 아님을 증명  
: 높은 음고를 가진 2개의 합성된 목소리(vibrato, non-vibrato)를 실험 대상에게 들려주었으나 모음의 인식에 있어서 큰 차이가 없음을 발견; 하지만 이 실험에서 사용된 비브라토는 매우 비인간적이었으며, 실험대상에게 합성음이 '사람 목소리'라고 가정하도록 요구했다. '사람 목소리' 같지 않은 소리를 들은 청자는 그 안에 '모음' 성분이 있음을 전혀 느끼지 못했던 것으로 보인다.
- John Chowning이 제시한 예는 Sundberg의 발견에 대치되는 사실을 발견  
: FM 음합성 방식을 통해 제작된 Chowning Tone은 무작위적인 성질이 추가된 주기적 비브라토를 적용하여 만든 합성된 목소리로, 모음의 인식을 보다 명확하게 만들어 주었음(Sundberg의 합성음은 여러 배음의 AM을 통해 생성)
- 합성음에 인간적인 비브라토, Drift, Jitter등의 요소를 추가함으로써, 청자로 하여금 소리를 '인간의 목소리'로 분류 및 인식하도록 할 수 있다. [Trk. 70]
- 잡음은 고음역에서의 모음 인식에 큰 영향을 끼침 [그림 9]



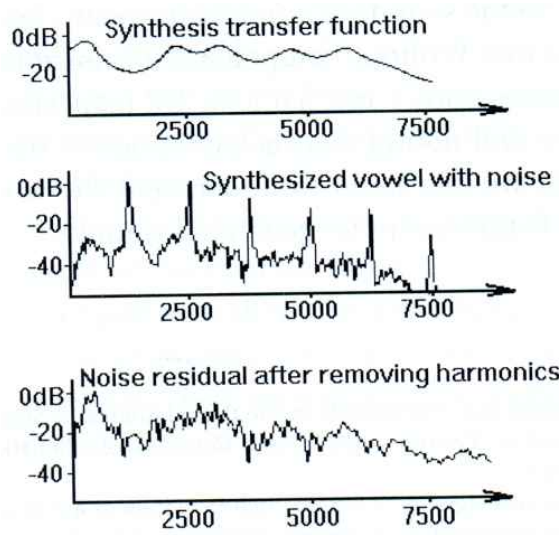


그림 9 잡음의 존재와 모음의 인식

- : 고음역에서는 배음 간의 간격이 넓어 실제의 음색형태를 판단하기 어려우나, 음색에 관련된 중요한 배음을 제거하고 남은 잡음의 음색형태를 살펴보면 원래 소리의 음색형태와 비슷함을 알 수 있음
- : 잔여 잡음성분(residual)만을 들어보면 소리를 구성하는 전체 성분에서 ‘속삭이는 소리처럼 들리는 부분’만 추출해낸 것처럼 들림

## 16.6 Conclusion

- 모든 자연음은 비주기적 성분을 갖고 있음
- 주기와 잡음의 편차는 사람의 목소리를 비롯하여, 관악기, 현악기 등 다른 악기에서도 발생
- 잡음과 비주기성 없이 합성된 소리는 기계적으로 들림
- 이러한 자연음의 성질을 면밀히 관찰하면, 더 자연스러운 합성음의 제작이 가능