

I . Objective Measures

수년간 잔향 시간이 강당의 음향적 행태의 유일한 관찰대상이었다. 심지어 오늘날에도 많은 건축가들에게 있어서도 마찬가지이다. 그러나 실내의 소리행태에 대한 더 많은 관점들이 있다. 이번 topic에서는 다음과 같은 추가적인 관찰대상을 다루고자 한다.

- 초기 감쇠시간
- 음의 맑음과 명료도
- 확산감(공간의 느낌 or 인상)
- 강연의 명료도(이해하기 쉬움)

잔향지역에 있어서 초기 작업은 사람들이 일부 초기 도착 소리에너지에 의해 가장 영향을 많이 받는다는 것을 보여주고 있다. 소리의 감지방향이 전적으로 초기전달임펄스(직접음)에 의해서 결정된다는 사실은 잘 알려져있다. 그래서 2차 임펄스(반사음)에 있어서 음원의 크기와 공간의 종류에 대한 더 많은 인지가 필요하다.¹⁾

또다른 보다 생리적인 영향은 음의 통합주기이다. 만약에 반사음이 직접음 뒤에 일정주기내에 도달하게되면, 양쪽귀에 직접음으로 통합될 것이다. 이런 방식으로, 매우 이른 반사는 실제적으로 직접음의 수음 레벨을 증가시켜서 signal-to-sound ratio를 강화시키게 된다.

1) Early Decay Time

앞서 논의되었던 것처럼, 잔향시간은 음원으로부터 소리가 확실히 멈춘 뒤에 잔향음이 60dB 감소할 때 까지 걸리는 시간을 말한다. 이상적인 공간에서의 이런 감쇠는 소리레벨에 대해 그래프화된/그려진 직선으로 나타나는 지수함수이다. 그러나 실제적인 강당의 연구에 따르면 항상 그런 것은 아니다.

(1973 Kuttruff의)연구에 따르면 후반부가 항상 새로운 소리에 의해서 **차폐**²⁾되기 때문에 잔향에 대한 주관적인 성향에 따라 초기 감쇠커브의 양상이 다르다. 이로인해서 초기감쇠시간 (**EDT:Early Decay Time**)³⁾이 사용된다. 이것은 작업에 따라서 감쇠의 초기 10~15dB이상만 일반 잔향시간과 같은 방법으로 측정된다.

2) Clarity⁴⁾ and Definition

음의 맑음과 명료도는 개별의 소리가 일반적인 가청 기류 내에서의 구분이 용이함을 말한다. 소리의 이런 흐름은 여러 가지의 형태로 나타날 수 있다 : 대화, 한 구절의 음악, 소리치는 경고, 기계의 진동 등 어느 것이든지 물론, 명료도의 정도는 내포된 특정음과 주로 관계가 있다. 그러나 건축적인 입장에서 보면, 명료도는 처음과 나중 도달 음의 에너지의 양사이의 비율

1) 청중이 소리 듣는 순서: 음원으로부터의 직접음→벽체 및 천장의 초기반사음+실의 다중반사음(잔향)

2) 차폐(masking 마스킹효과) 큰소리와 작은 소리가 동시에 존재하면 큰소리로 인해 작은 소리가 들리지 않는 현상

3) Early Decay Time (EDT, 초기 감쇠 시간)

: 잔향 시간의 부가 파라미터로서, 음원이 정지된 후 10dB 감쇠하는데 걸리는 시간이다.

RT에 비해 소리의 주관적인 울림과 잘 일치한다고 한다.

완전 확산 음장의 경우에는 RT와 일치하지만, 일반적으로 RT보다 적게 나타난다.

4) Clarity (C80) : 초기 에너지 대 후기 에너지(early/late sound ratio)의 비율로서 음악의 명료도를 나타낸다.

에 따라 다르다. 초기음의 명료도는 절대적이지 않기 때문에, 이러한 비율이 존재하는 것이다. 그러나 어떤 형태에서든지, 이러한 측정은 벽에 의해서 음악적으로 나타나는 ‘섞이는’ 양의 평가를 가능하게 한다.

개요에서 언급된 바에 따라, 매우 유사한 두개의 음이 (50~80msec의) 아주 근소한 차이의 시차를 가지고 귀에 도달했을 때, 귀는 한 음원의 소리로 합쳐져서 들릴 것이다. 이것은 통합 주기내에 도달하는 반사음의 에너지는 직접음의 수신강도를 효과적으로 증가시키는 것을 의미한다. 이것은 상당히 유용한 정보중 하나로 뜻하는 바와 같이 음향학자들은 근접 반사를 이용하여 낮은 소리 레벨과 관련하여 발행하는 많은 문제점들을 해결하기 위한 방안들을 갖게 된다.

-Deutlichkeit(D):

지금까지 ‘명료도’로 간주되어온 것은 강연장 법규의 적용을 위해서 Theile에 의해서 제안되었고, 초기 50msec이내의 음에너지와 모든 도달음에너지의 총합 사이의 선형적인 비율이다.

-Klarheitsmass(C80):

지금까지 ‘명료도’로 간주되어온 것으로 음악에 더 적당한 것처럼 Reichardt에 의해서 제안되었다. 초기 80msec 이내와 나중에 도달된 음사이의 에너지를 측정하는 로그비이다.

이런 수치상 최적의 값을 고려할 때, 소리의 특별한 형태나 음악스타일은 매우 중요하다. 예를들어, 솔로 기타리스트는 모든 관객들이 잘 분간해서 들을 수 있게 연주해야한다. 반면에 현악의 소리는 작은 잔향이 있는 것이 매우 좋다. 이러한 의문을 고려할 때 다음의 논리적 근거에 의해서 용이하게 증명될 것이다.

음악적 경험의 관점에서, 두 개의 매우 극단적인 현상이 일어날 수 있다 ; 실내악에 대한 깊은 이해는 바그너 교향곡 전체를 깨우치는 것에 비교된다. 각각은 청중의 올바른 이해를 위해서 다소 다른 음향 환경을 필요로 한다.

실내악 감상은 대부분 청중과 연주자사이의 근접성을 내포한다. 근접성의 주요한 특징중에 하나는 직접음장이 잔향장을 지배하기 위해서 나타난다는 것이다. 따라서, 친밀감을 증진시키는 환경은 통합주기 내에서 도달하는 처음과 두 번째 반사의 횟수는 최대화하는 반면에 긴 잔향주기를 최소화하는 것이다. 이러한 반사는 굳이 잔향으로 받아들여지지 않고 직접음을 더욱 강화시키는 작용을 한다. 이러한 접근의 결과는 초기 35~50msec 이내의 초기 음에너지를 최대화하는 반면에 나중 음에너지는 최소화한다. 반면에 영감을 받거나 감동시키는 극단적인 상태는 보다 넓은 거리감을 요구할 것이다. 이런 방식에서 청중은 정말로 간접적인 현상을 지각하고 있다는 것을 느낄 수 있다. 이런 경우에, 적정 잔향을 강화함은 물론 음원의 세기를 강화하기위해서 초기반사를 최대화해야한다.

이런 효과를 더욱 증대시키기 위해서 역제곱법칙에 의하면 잔향은 증가하더라도 직접음은 감소되므로 이미 음원에 가까이 있는 청중에게 도달하는 초기 반사를 최소화하는 것이다. 그 래야 음원으로부터의 거리가 다양한 청중에게 초기 도달음의 양이 ‘잘 조율되는’ 건축물이 될 수 있다. 대규모의 웅장한 오페라의 경우 매우 적은 초기 에너지는 매우 둔탁하고 불분명한 소리를 생성함에 따라 이런 경우의 목적은 비율을 줄이는 것이 아니다.

3) Spatial Impression(확산감)⁵⁾

이런 효과를 더욱 최대화하는 것은 역제곱법칙에 의하면 잔향은 증가하더라도 직접음은 감소되므로 이미 음원에 가까이 있는 청중에게 도달하는 초기 반사를 최소화하는 것이다. 그래야 음원으로부터의 거리가 다양한 청중에게 초기 도달음의 양이 '잘 조율되는' 건축물이 될 수 있다. 대규모의 웅장한 오페라의 경우 매우 적은 초기 에너지는 매우 둔탁하고 불분명한 소리를 생성함에 따라 이런 경우의 목적은 비율을 줄이는 것이 아니다.

따라서 확산감의 측정은 직접음과 관련이 있는 각 반사각에 기초를 두고 계산된다. 이런 신체적수치의 첫 번째인 **측음 에너지 비율(Lf:Lateral Energy Fraction)⁶⁾**은 초기 80msec 이내에 총에너지가 측음 에너지의 선형적인 비율로 Barron & Marshall에 의해서 제안되었다.

???

여기에서:

Q=반사경로와 두 귀의 중심선 사이의 각
 $\cos Q = |\sin A| \cos B$ 로 대체될 수 있다.

여기서, A= 상대적인 방위각

B= 상대적인 위도

이러한 관계로부터, 청중은 음원에 집적적으로 직면해 있다는 가정하에, 공간감을 위한 부분적인 반사의 효과는 A가 90도로 가까워지고 B가 0도로 가까워질 때 공간감이 최대로 될 것이다. 이것은 소리가 가까운 쪽 귀에 직접 전달되어 최대 내부 청각의 지연과 머리 섀도잉 효과가 발생하는 것을 뜻한다. 대신에 확실한 값은 머리 양 단부에 도달하는 반사는 동일하게 측정되도록 간단히 제안된다. 시행된 임펄스 모델의 해석에 의해, 측음에너지의 비율은 쉽게 결정될 수 있다. 음의 밝음과 명료도의 수치와 같이 주어진 Lf 는 비율을 나타내고, **임펄스⁷⁾**를 정의 하기위해서 사용된 음의 세기 단위는 근거가 없는 것이다. 따라서, 앞뒤가 맞지 않는 에너지 증가의 가정하에 제곱 **임펄스 반응(Impulse Response)⁸⁾**이 사용되었다. 이 수치로부터, 공간감의 정도는 다음의 제안된 공식으로부터 추정할 수 있다.

$$SI=14.5(Lf-0.05)$$

첫 번째 공식으로부터, 개별적인 공간감의 성향을 고려할 때 각 반사의 지연은 상대적으로 5msec과 80msec사이의 상대적으로 중요하지 않은 값으로 나타낸다. 5msec 이하의 지연에 있어서, Barron & Marshall은 마스킹효과가 어떤 확산감의 요소를 상당히 감소시킨다고 추론하였기 때문에 이러한 반사는 총 음에너지에만 포함된다.

또다른 수치인 상관관계 정도(K)는 Gottlob에 의해서 제안되었다. 이것은 측음 에너지의 측정값이 아니라 비 측음에너지의 측정값이다. 정의된 것처럼, 비 측음에너지는 중간층으로부터

5) Spatial Impression(확산감) : 두귀효과에 의해 얻어진, 음장의 공간적 인상의 하나.

6) Lateral Energy Fraction (LEF, **옆방향 에너지 비율**)

측면 방향으로부터 도달되는 초기 반사음 에너지로 확산감에 큰 기여를 한다.

그리고 모든 방향으로부터 도달되는 직접음 및 초기 반사음의 에너지 사이의 비율(%)이다.

7) Impulse (**임펄스**) : 펄스 중에서 변화가 급격하고, 지속 시간이 충분히 짧은 펄스를 말한다.

이 속에는 넓이가 유한이고, 지속 시간이 무한히 작아진 극한의 이상화된 펄스도 포함된다.

cf) Pulse (**펄스**) : 어떤 정상 상태에서 진폭이 천이(shift)하여 유한시간만큼 지속하고, 원래의 상태로 돌아가는 파형.

8) Impulse Response (**임펄스 응답**): 시각 t=0에 대해서 단위 임펄스를 입력했을 때의 시스템의 시간 응답.

±60도이내에 도달하는 반사의 결과로 간주된다. 이 값은 다음과 같이 나타낼수 있다.

???

임펄스 모델을 사용하여 0.1미만으로 지연시키거나 $-60 < a < +60$ 범위로 2개의 Boolean test 를 실시한 각 임펄스로 L_f 와 같은 방법으로 계산 할 수 있다.

4) Speech Intelligibility⁹⁾ (강연의 요해도 ,즉 이해하기 쉬움)

오늘날의 멀티미디어 사회에서 조차도 개인적인 의사소통의 차원에 있어서, 강연은 가장 중요하고 효과적인 방법 중에 하나이다. Chapanis에 의해 행해진 실험에 의하면 예상했던대로 강연이 허용되지 않았을 때와 비교해서 그룹 내에서 수행되는 협동작용의 수행 시간이 10분가량 더 빨랐다. 따라서, 개인과 그룹사이에서 전적으로 강연을 위해서 사용된 많은 방은 음향적인 설계를 고려하고 그런 사용을 강화하는 것이 중요하다.

① Measuring Intelligibility(요해도의 측정)

강연의 명료도는 일반 청중이 강연에 사용되는 단어나 문구의 정확한 이해의 정도를 나타낸다. 강연을 통해서 언급되는 정보는 문맥상, 시각적으로, 체스처의 범주내에 포함된 것이므로, 심지어 개별 강연의 일부분만을 정확하게 듣는다고 하더라도 강연의 의미를 이해하는 것은 가능하다. 그러나, 대규모 강당과 대형 공간에서는 강연이 재생되므로 청중은 이러한 연설에 접근은 제한되어있고, 실제적으로 입에 의해서 생성되는 소리에 보다 신중하게 접근해야한다.

이런 지역에 대한 연구는 20세기 초 전화와 텔레커뮤니케이션 시스템의 발달로 시작되었다. 이 연구의 성과는 분명한 발음 테스트에 기초한 명료도를 위한 음의 측정이었다. (Lochner와 Burger가 묘사한 것처럼) 이 과정은 일반적으로 실험실내 한 명이상의 청자에게 단어나 문장의 음절 단위로 아나운서가 낭독하는 것으로 구성된다. 청중에의해서 이것이 정확하게 녹음될 퍼센트를 articulation score라고 일컫는다. 그리고 실험실에서의 음성 명료도의 측정 수치를 'in-situ'로 사용한다.

명료도 테스트 방법은 실제적으로 Bell Telephone Laboratories에서 정제되었고, 후에 하버드 대학의 Psycho-Acoustic Laboratory (신체 음향 연구실)에서 이루어졌다. 후기 작업으로부터, 음성적으로 균형이 잡힌 세트, 단음절의 테스트 명단이 준비되었고, 음성적으로 균형이 잡힌 세트는 하버드 P.B.50단어 스코어를 말한다. 측정되는 명료도의 비음성적인 영향을 부인하기 위해서, 이런 단어들은 의미가 없거나 아무렇게나 섞어 놓은 음절만으로 구성되어있다. 그래서 청자에 의해서 정확하게 녹음하기 위해서는 각각의 자음과 모음이 명확하게 들려야한다. 더 정확한 측정을 위해서는, 강연을 유창하게 하기위해서 실제로 사용되는 표현을 삼입한 음절로 많은 실험들이 수행되어야 한다. 현재 이런 방법론의 많은 기원이 있다.(Bradley 와 Latham에 의해서 사용된 Fairbank 리듬법과 같은) 그러나, 유효 수치는 정확하게 녹음되는 음절의 퍼센트이다. 그래서 명료도의 정도는 이런 수치의 평균과 상관관계가 있는 것으로 간주된다. 이런 퍼센트는 부분적인 울타리의 비율로 음성의 명료도를 측정하는 것이다.

9) Intelligibility(요해도) : 의미 있는 단어나 문장을 단위로 구성된 시험 음성을 듣고 정확하게 들은 것의 백분율. 단어를 단위로 한 경우의 요해도를 단어 요해도, 문장을 단위로 한 경우의 요해도를 문장 요해도라고 한다. 요해도 값은 시험 음성의 구성법에 의해 다르므로, 이용한 리스트를 명기할 필요가 있다.

앞서 언급된 것처럼, 보통 연계된 강연은 음절의 일부가 잘 안 들리더라도 이해할 수 있다. 이는 청중이 문장의 문맥으로부터 의미를 추론할 수 있기 때문이다. 그러나, 완벽한 조건하에서조차, 일반적으로 도달할 수 있는 최대 단어 점수는 불가피한 예러 때문에 95%정도가 된다. 단어 스코어 80%는 청중이 굳이 들으려고 노력하지 않아도 모든 문장을 이해할 수 있는 수치이다. 단어 스코어가 70%정도 되는 실에서는 청중이 강연을 이해하기 위해서 집중을 해야한다. 반면에 60% 이하는 명료도가 상당히 불량한 상태이다.

② Predicting Intelligibility(요해도의 예측)

실내에서는 강연의 명료도를 예상할 수 있는 몇 가지 방법이 있다. 이것은 **명료도지수(AI)**¹⁾ **대화간섭레벨(SIL)**¹⁾, 가중 신호대잡음비(A-weighted signal-to-noise ratio ; Lsa), 사용가능한 음과 해로운 음의 비율(useful/detrimental sound ratio's: U80과 U95) 그리고 **음성 전달지수(speech transmission index : STI)** 이러한 각 방법들은 동일한 원리에 기초를 두고 있고, 강연 수음자와 간섭 소음레벨사이의 비율을 결정한다. 강연의 요해도를 결정하는 것은 기본적인 신호대잡음의 관계이다.-비율이 더 높을수록 요해도도 높아진다.

이번장의 취지를 위해서, 강연은 특별한 강연 단위의 정확한 해석을 위해서 인식이 가능한 음성 정보로만 간주되어야한다. 청중에게 전달되는 모든 다른 음에너지는 간섭 잡음으로 간주되어야한다. 따라서 신호대잡음비는 동일한 신호로 그러나 주파수 밴드는 다른 범위에서 측정될 수 있다. 이런 신호대잡음비에 영향을 주는 기본적으로 측정 가능한 3요소가 있다:

- 강연 발표 레벨과 매너
- 암소음의 레벨과 스펙트럼
- 실 반응의 특성과 기간

③ Speech Level (음성레벨)

교란이 없는 환경에서는, 일반적인 음성레벨이 55dB과 65dB사이로 떨어진다. (화자로부터 1m떨어진 거리에서 측정시) 특별한 경우에, 주의를 주기위해 소리를 지를 때는 레벨은 96dB까지 높아질 수 있고, 가볍게 속삭일 때는 30dB까지 낮아진다. 음성레벨은 음향적인 조건이 같더라도 개인별로 매우 다양하다는 것을 명심해야한다. 그래서, 강연의 명료도를 예상할때, 평균치보다는 최악의 경우의 수치를 기본값으로 설정한다. 이것은 내부 개인 강연레벨의 표준편차를 나타내는 양에 의해서 평균치를 조합함으로써 행해질 수 있다. 극단적인 노력으로도 변수를 완벽하게 설명할 수 없다. 여성의 15%가 75dB이상으로 음성이 올라갈수 없는데에 비해 남성이 소리를 지르면 15%는 96dB를 쉽게 넘고, 종종 104dB까지 이르기도 한다. 다음의 표는 vocal effort를 위한 '최악의 수치'를 나타낸다.

10) Articulation Index (AI) : 명료지수

회화 방해(Speech Interference) 정도를 표현하는 방법.

명료지표를 얻는 방법은 말하는 사람과 듣는 사람 사이의 신호와 잡음의 차이에 가중치를 1/3옥타브 대역별로 곱하고, 이들을 합한값으로 얻게되며 시가 0%이면 대화가 불가능한 것으로 시가 100%이면 완벽한 회화가 가능한 것을 의미하게 된다.

11) Speech Interference Level(SIL) 대화 간섭 레벨

대화를 나누는데 있어서 암소음의 영향 고려시 이를 평가를 위한 수치.

Vocal effort	dB(A)	Vocal effort	dB(A)
Whispering	32	Raised	57
Soft	37	Loud	62
Relaxed	42	Very Loud	67
Normal(private)	47	Shouting	72
Normal(public)	52	Max. shout	77

-표 1. 평균 vocal effort와 소리레벨 표-

명료도를 계산할 때, 청중의 귀의 레벨은 상관없이 큰소리로 말하거나 소리를 지르는 강연이 더 이해하기 힘들다는 것을 알았다. 이것은 주로 75dB이상에서는 현저히 나타나는 음성과 역량의 변화때문이다. 게다가 만약에 보다 노멀한 강연의 경우, 귀에 전달되는 레벨이(80dB이상) 매우 높다면, 귀의 과부하를 방지하기위한 연구가 행해질 것이다. 이러한 현상은 강연자와 청중의 거리가 매우 가까워서 발생하게 된다.

계산을 간단하게 하기위해서, 이러한 현상은 다음과 같은 규칙으로 정리될수 있다. 우선, 큰 소리로 연설을 할경우, 10dB 마다 출력 레벨은 75dB이상 상승한다.(발표자로부터 1m거리에서 측정한 경우), 수신되는 강연과 간섭 잡음사이의 신호대잡음비는 4dB로 감소되어야한다. 두 번째로, 일반적인 강연의 경우, 45dB과 75dB사이의 출력레벨과 80dB이하의 수신레벨에서는 눈에 띄는 현상이 없다고 가정하였다. 극단적인 근접효과는 80dB이상의 경우 매 10dB마다 명료도는 거의 3~5%정도 감소하고 신호대잡음비는(간섭잡음의 적당한 낮은 레벨로 가정한다.) 1dB보다 조금 감소하는 것으로 나타났다.

④ Background Noise(암소음)¹²⁾

모든 음향적인 환경에서는, 항상 당면하는 암소음의 일정한 레벨이 있다. 이 레벨은 대부분 공간과 보다 즉흥적인 상황에서 발생하는 활동에 따라 다르다. 암소음의 가장 눈에 띄는 현상은 음성신호를 차폐한다는 것이다. 그래서 수음자가 강연에 특별히 집중을 함에 따라 신호대잡음비가 줄어든다. 그러나 암소음의 다른 효과도 있다.

단기 강연은 중요한 vocal 효과를 수반하는데 반해서 장시간의 대화는 화자에 의해 안정적으로 유지될 수 있도록 훨씬 더 낮은 레벨을 요구한다. 그러나, 이것은 화자의 강연레벨 선택의 관점에만 해당된다.

음성레벨의 가장 중요한 결정요소 중의 하나는 Lombard 효과가 나타나는 것이다. (Lombard에 의해 제안되었고, Lane과 Tranel이 더 자세히 연구한)이 효과는 헤드폰을 끼고 들으면서 말할때 가장 확연하게 나타난다. 일반적으로 강연자는 자신의 히어링으로 피드백을 거쳐 자기 자신의 vocal effect를 체크한다. 헤드폰을 낀 채로 음악을 따라서 흥얼거리면, 목소리는 거의 자동적으로 커지게 마련이다. Lane과 Tranel에 따르면 암소음과 다른 간섭잡음은 일시적이기도 하고 영구 청력손실을 유발하는 등 매우 유사한 현상을 보인다고 한다.

이런 효과를 정량화하는 것은 개인적인 반응을 복합하는 것처럼 어렵다. 예를 들어, 일상적인 대화에서, 사람들은 목소리를 높이기보다는 더 가까이 모여서 대화를 한다. 그러나

12) Background Noise(암소음) :

어떤 음을 대상으로 생각할 때 그 음이 아니면서 그 장소에 있는 소음을 대상으로 하는 암소음이라 한다.

(Lazarus에 의해 제안된) 엄지손가락법칙(rule of thumb relationship)의 관계는 45dB이상의 간섭잡음이 1dB증가할 때마다 출력 강연레벨은 평균적으로 0.5~0.6dB정도 상승할 것이다. 이런 자동적인 상승은 얼굴을 맞대고 하는 개인적인 대화처럼 약한 스피치레벨에서는 좀처럼 일어나지 않는다. 이런 상황에서는, 대부분 더욱 선명한 대화를 위해서 화자들이 더욱 가까이 접근을 하게 되고 위에서 언급한 효과는 55dB이상의 스피치레벨을 필요로 하는 경우에만 해당된다.

⑤ Room Response (실 응답)¹³⁾

실의 반응 특징은 강연의 명료도에 영향을 줄 수 있다. 이로울 수도 있고 해로울 수도 있는 이 효과는 임펄스 반응 함수이다. 일반적으로 강연의 통합주기(35~50msec)이내에 청중에게 도달되는 에너지의 양이 상대적으로 높을 때 울타리는 강연의 인지를 강화시킬 것이다. 주어진 암소음은 지속적으로 울타리에 반사되고 추가적인 초기 반사는 신호대잡음비를 효과적으로 상승시킬 것이다. 그러나 나중에 도달하는 반사음과 과도한 잔향은 직접 스피치신호와 간섭을 이루면서 암소음레벨을 사실상 증가시키게 된다. 따라서, 너무 늦게 도달하는 음에너지는 신호대잡음비를 감소시키는 경향이 있다.

5) Measures of Speech Intelligibility (요해도의 측정)

① The A-Weighted Signal/Noise Ratio(SNA)¹⁴⁾

이것은 지금까지 제안된 모든 방법들 중에서 수행하기 가장 쉽고 간단하다. 결론만 간략하게 말하자면, 이 측정치는 특정시간에 측정된 The A-Weighted long-term의 평균 스피치레벨과 The A-Weighted long-term의 암소음의 평균 레벨사이의 차이와 관계가 있다.

$$SNA=LSA-LNA$$

② The Articulation Index (AI) [명료도지수]

이 수치는 기본적으로 (0.25, 0.5, 1, 2, 4kHz를 중심주파수로 하는)5옥타브 밴드¹⁵⁾에서 신호대잡음비의 계산에 기초해서 0.0에서부터 1.0의 범위에서 직선적으로 측정되었다. 더욱 정

13) Room Response (RR, 룸 응답)

: 공간적 인상을 나타내는 목적으로, Jordan에 의해 제안된 음향 지표의 하나이다.

0~80ms까지의 전체 에너지에 대한 확산감을 주는 25~80ms의 죽음에너지와 80~160ms의 잔향 에너지 비를 표현한다.

14) signal to noise ratio (신호대잡음비)

: 신호의 품위 레벨의 척도로 원하는 신호(Signal)와 원하지 않는 신호 혹은 잡음(Noise)의 비로서, 보통 S/N비라고 표기한다. 따라서 S/N비가 클수록 양호한 신호라고 할 수 있다.

S/N비는 일반적으로 dB로 표시하며, S/N비를 높이기 위해 다수의 신호를 채취하여 평균을 구한다..

신호전력을 S, 잡음전력을 N이라 할 때 $10 \log_{10}(S/N)$ 으로 나타낸다.

15) Octave Band (옥타브 밴드)

: 상한 주파수 f2와 하한 주파수 f1의 비가 2가 되도록 분할한 주파수 대역.

주파수를 대수축 상에 놓을 경우, 옥타브 밴드는 등 간격이지만, 그것을 다시 1/n마다 분할하면 각각의 주파수 대역을 1/n 옥타브 밴드라고 한다.

cf) Octave (옥타브)

1옥타브는 최고 주파수가 최저 주파수의 두배가 되는 주파수 대역을 말한다.

예를들면,40Hz의 1옥타브는 80Hz와 20Hz이다.

밀한 데이터가 필요한 경우 1/3 옥타브에서 주파수분석을 한다. 그러나 이 경우 스피치 스펙트럼과 소음 스펙트럼의 보다 자세한 정보가 있어야 한다. 일반적으로 스피치레벨이 노말한 화자를 위한 long term value에 따르므로, 옥타브 스펙트럼은 간단한 계산에 적절하다.

AI 계산의 3단계

- 각 옥타브 밴드를 위한 유효 신호대잡음비 측정
- 각 비율에 가중요소를 적용하고 최대 +18dB, 최소 -12dB을 확실히 고정하라.
- 평균치 계산

따라서 분절음의 색인은 다음과 같이 계산될 수 있다.

$$AI = \frac{G_{ij}}{30dB} \sum_{i=1}^5 (L_{sa} - L_{na} + 12) \text{ dB}$$

여기에서, $G_{[i]}$ = 각 옥타브 밴드의 가중요소

③ Speech Transmission Index (STI) [음성 전달지수]¹⁶⁾

Houtgast와 Steeneken에 의해서 강연의 명료도의 측정수단으로 처음 제시된 음성전달지수의 유도는 기본적으로 명료도지수(the articulation index :AI)보다 훨씬 더 자세하다. 더 중요한 개선점 중의 하나는 시간적인 영역에서 소리의 일그러짐을 포함하도록 시도한 것이다.

앞서 논의된 것처럼, 이러한 소리의 일그러짐은 잔향과 지연된 반사로 인해서 야기된다. 이런 현상의 원리는 암소음의 효과적인 간섭레벨을 상승시키는 2차 효과를 가지는 스피치 신호의 강도에 있어서 빠른 파동의 주름이 퍼지는 경향이 있다. 그래서, 음성을 강도와 스펙트럼 상에서의 일시적인 변동이 있는 음에너지의 흐름으로 간주하여 음성 전달에 관한 법규를 제정하였다. 따라서 파동이 전달 시스템에 의해서 보존되는 정도는 정확성의 측정치로 간주되었다.

Houtgast와 Steeneken은 음성신호의 일시적 기밀의 보존이 음성의 개별적인 사인곡선의 요소를 내포하고 있음을 논의했다. 이런 이유로 그들은 화자에 의해서 만들어져 수음자에게도 존재하는 사인 음파를 연장함으로써 실의 음향적인 장점을 결정하도록 제안했다.

그래서, 장시간 연설의 평균 음 레벨은 일상의 대화와 유사한 스펙트럼을 갖는 이론적인 모의 신호로 대체된다. 이 신호의 강도는 1의 변조지수¹⁷⁾를 가지고 있는 함수에 의해서 조절된다. 이런 방식으로, 어떠한 신호의 감쇠는 수신신호로부터 유추한 변조지수내의 감소로 나타날 것이다.

일반적인 울타리에 있어서, 감쇠 잔향의 비율은 상대적으로 안정적으로 주어져, 신호의 더

16) Speech Transmission Index (STI) [음성 전달지수]

: 음성전송계의 물리적 바로미터로 명료도를 추정하는 것이다.

각 음성 주파수대역별로 구해지는 변조주파수에 있어서의 변조지수(MTF :Modulation Transfer Function)를 바탕으로 인간의 실험감각을 가미하여 구한다.

- 0에서 1사이의 수치로 표시된다. *일반적으로 STI가 0.6이상이면 좋다고 되어있다.

17) Modulation (변조) 일정한 고주파 전류 또는 펄스 전류에 음성과 영상신호 등으로 얼마간의 변화를 주는 것.

느린 변화에 대한 효과는 더 빠른 변화와는 다를 것이다. 일상적인 대화에 대한 연구에서는 1Hz와 8Hz사이의 변동주파수는 3Hz에서 최고로 나타날 것이다. 그래서 Houtgast와 Steeneken은 매우 느린 음성과 매우 빠른 음성 모두를 아주 잘 이해하기 위해서는 0.4Hz와 20Hz사이에서 정확한 전송이 이루어져야한다고 제안한다.

만약에 F= 변동주파수를, t=시간(초)를 나타내면, 모의 신호의 강도는 다음의 함수에 따라서 변한다 : $1+\cos(2\pi Ft)$

이것을 임펄스모델과 연관하면, 다음의 관계가 변화 인덱스를 위해 주어진다.

$$m(F) = \frac{\sum_n \frac{a_n \exp(-j\pi F \frac{r_n}{c})}{r_n^2}}{\sum_n \frac{a_n}{r_n^2}}$$

여기에서,

c= 개별적인 감쇠에 작용하는 요소와 소리의 속도

r_n= n번째 임펄스의 상대적인 경로의 길이

임펄스의 반응은 각 F값을 공식에 18번 적용한 것이다. 변조지수 8의 경우, 울타리의 변동전달함수(Modulation Transfer)로 나타낼 수 있다. 그러나 최근에는 암소음의 효과를 고려하지 않고 계산한다. 암소음은 변동 인덱스의 계산을 수신자의 위치에서 측정된 수신 신호의 평균 강도로 조정함으로써 포함될 수 있다.

$$m'(F) = m(F) \frac{I_s \sum_n \frac{a_n}{r_n^2}}{I_s \sum_n \frac{a_n}{r_n^2} + \epsilon}$$

이 공식 내에서, 잔향과 직접음장의 효과가 모두 나타난 반면에 적절한 신호대잡음비로 쉽게 전환할 수 있다 :

$$SNR = 10 \log \left(\frac{m(F)}{1 - m(F)} \right)$$

각 개별 비율을 +15dB과 -15dB사이의 값으로 제한한 뒤에 잡음비의 평균 과다 신호는 18이라는 같은 가중치로 전환된다. 그래서 STI는 다음의 식에서처럼 선형적인 증가율을 보인다.

$$STI = \left(\frac{SNR + 15}{30dB} \right)$$

④ Useful/Detrimental Ratio's(사용가능한 음에대한 해로운 음의 비율)

이러한 측정치는 초기에 논의 되었던 것과 유사한 나중 도달음에 추가되는 암소음에너지의 효과를 가진 초기에대한 후기 음의 비율이다. Lochner와 Burger는 초기95msec안에 도달하는 음에너지의 가중치의 합으로 계산된 사용가능한 에너지를 음의 명료도의 기준척도로서 비율의 개념을 처음 도입했다. 그래서 해로운 음은 암음이 추가된상태로 나중에 도달하는 음에너지이

다. 사용가능한/해로운 음에너지를 추론하기위해서, 초기에 대한 나중 에너지의 비율이 우선적으로 정해져야 한다. Lochner와 Burger 식은 다음의 방정식에의해 주어진 C95와 관계가있다.

$$C_{[95]}=10\log$$

여기에서 m=각 개별 반사가 유용한 에너지의 합으로 간섭된 에너지의 파동

m은 주관적인 관측 한계내에 있기때문에 계산이 다소 막연하다. 기본 가설은 95msec보다 더 이른 특별한 임펄스에 대한 것이라는 것이다. 이 음에너지의 일부는 직접음과 합쳐질 것이다. 일부분은 상대적인 지연과 상대적인 레벨의 함수로 바뀐다. m값의 근사치는 Bradley에 의해서 다음과 같이 주어졌다.

$$m=(2.3-0.6a^{0.7})+(0.0248t-0.00177a^{1.35}t)$$

여기에서,

a= 반사의 상대적인 진폭

t= 상대적인 지연 시간

이 초기/나중비로부터, 사용가능한 음과 해로운 음의 비(U95)는 다음의 조건에서 구할수 있다.

$$U_{95}=\frac{C[95]}{1+(C[95]+1)(EBL/ESL)}$$

여기에서,

E_{BL} 과 E_{SL} =주변과 음성의 총 에너지로,

$E_{BL}=10(BL/10)$, $E_{SL}=10(SL/10)$

SL과 BL=장기간 안정상태에서의 암음과 음성레벨의 제곱의 평균(rms:root mean square)

위의 방정식을 사용해서 Bradley는 다른 초기에대한 나중비가 유사하게 적용될수 있다고 제안했다. 예를들어, U80은 C80으로부터 유추할 수 있다. 이러한 수치의 이점은 Lochner와 Burger에의해서 정의된것 처럼 복잡한 가중 과정을 요구하지않기 때문에 더 간단하게 계산할 수 있다는 것이다. U95와 C80으로 측정된 음성의 명료도로부터 Bradley는 각 그룹의 데이터에 가장 적합한 3차다항식으로부터 다음의 관계를 유추하였다.

$$SI=0.7348U_{95}-0.09943.U_{95}^2+0.0005457.U_{95}^3+197.39$$

$$SI=.219U_{80}-0.02466.U_{80}^2+0.00295.U_{80}^3+95.65$$

이런 그리고 다른 측정치의 상대적인 장점이 연구되고부터 Bradley는 U80이 더욱 안전하고 일반적으로 명료도의 예측의 신빙성이 더 있다고 제안했다.