

기술해설

음성품질 주관 평가법
Subjective Assessment of Speech Quality

김정환·김민택·강성훈*

(Jeong Whan Kim, Min Taek Kim, Seong Hoon Kang)

*연세대학교통신연구소 연구책임연구원

I. 서론

다음의 일일 생활자를 대상으로 하는 전화전송분야의 음성품질의 좋고 나쁨은 가장 중요하고 고려하여야 할 것의 하나이다. 많은 연구자들이 각종 전송수단을 구별하고 전송기준을 설정하려는 것은 가장 경제적인 비용의 음성품질 실현에 그 목적이 있다. 그러나 이러한 음성품질의 좋은 정도(goodness)는 동일한 전송매체의 독립적인 판단으로 결정된다. 즉 다른 나라에서 사용하고 있는 전송기준은 그대로 보았더라도, 각 경우에는, 사회·문화적인 환경에 의한 품질의 요구 수준, 선비, 인명법 등 개인적 요인에 따라 많은 차이가 있기 때문에 국내 상황에 맞는 음성품질의 평가를 기대하기는 어렵다.

따라서, 지금의 통신환경에 맞는 또 일반 생활자들의 개인적 요인과의 전화전송분야의 물리적 파라미터들과의 함수관계에 대한 광범위한 독자적 연구가 이루어져야 한다. 이러한 이유로 해서 CCITT 제11, 12, 13회의 물리적 기준을 고려하여 음성품질에 대한 표준화된 전송기준 자료를 발표하지 못하는 실정이다.

본 논문에서는 전화 전송분야의 음성품질의 주관적 평가법을 소개하고, 이를 이용하여 음성품질의 주관적 평가를 수행하여, 전송매체와 전송수단의 선택에 있어 어떤 기준을 채택하여야 하는지를 연구하여, 이를 실제적 임의관정을 통해 평가

된 것이다. 여기서 앞으로 주관적이 결정기준이라고 한다면, 오히려의 평가¹⁾는 관능평가(sensory evaluation)라고 할 수 있다.

일반 생활자를 대상으로 하는 음성품질 주관 평가법은 직감과 사람에 의해 정해진 가변적이기 때문에 시간 가변적(time variant), 개인 가변적(personal variant)인 어떤 정량적 표현으로 나타내게 된다. 일단, 어떤 평가자가 구해지면, 그 음성품질의 측정치를 표현하는데 필요한 기술 방법적인 방법을 구하여야 한다. 그러나, 이렇게 편리한 주관 평가방법은 항상 시간 가변적, 개인 가변적인 성질에서 일어나는 것은 아니다. 이 성질들을 어떻게든 시간 불변적, 개인 불변적인 주관적 반응과 연관시키고, 이것을 매개변수로 하여 지능화시키는 물리적 측정과 관련되어야 한다. 즉, 음성품질 척도로는 통신키지를 선택하기 위한 파라미터로 변환하여 연산할 수 있어야 한다.

반응을 평가하는데는 그 대상과 목적에 가장 적절한 평가방법을 찾는 것이 중요하다. 문제의 첫수준에 따라 어떤 척도를 사용하는가에 따라, 그 목적으로부터 유도된 최적척도가 필요하다. 그러나 하나의 평가척도만으로 목적에 맞는 평가는 안될 경우는 많다. 따라서, 주관적 음성품질의 평가는 여러 척도를 사용하여 여러 척도를 사용하여 목적에 맞는 평가가 이루어지도록 해야 한다. 즉, 주관적 평가는 여러 척도를 사용하여 이루어지도록 해야 한다.

본 자료는 국내에서 음성품질 주관 평가법에 관한

*본 논문은 1987년 11월 10일 접수된 원고로, 1988년 1월 15일 게재된 것임.
*연세대학교통신연구소 연구책임연구원

연구를 체계적으로 수행할 수 있도록 하기 위해, CCITT의 연구결과를 비롯하여 각국의 오피니언 평가법을 체계적으로 정리하여, 국내 표준으로 활용하고자 작성한 것이다.

II. 종 류

1. 회화 오피니언 평가, 청취 오피니언 평가 및 송화 오피니언 평가

회화 오피니언 평가는, 어떤 특성을 가진 전화 연결망을 구성하고 한쌍의 피험자(2명)에게 통화로를 통해 회화를 하도록 한다. 피험자들이 통화로에 익숙해질 때까지 충분히 연습을 시킨 후에, 오피니언 시험 절차에 따라 품질척도 반응지에 통화품질을 평가하도록 한다. 회화 시험(conversation test)은 실험실(laboratory test)이나 현장(field test)에서 수행하며, 일반적으로 손실 효과(loss effect)를 평가하는데 이용된다.⁶⁾ 그러나 실험내용이 손실에 의존적일 때는 청취 오피니언시험을 수행하는 것이 적절하다.

청취 오피니언 시험(listening test)에서, 음성 시료는 미리 녹음하여 특정 레벨로 설정된 시험회선에서 재생하며, 청취자에게는 음성 샘플을 듣고 전체적인 통화품질에 대해 오피니언 평가를 하도록 한다. 그러나 청취 오피니언 시험은 정상적인 시간개념과 상호작용을 일으키는 지연시간, 에코 또는 어떤 다른 손실효과를 평가하는데 사용할 수 없다. 또한, 열화(degradation)의 양이 실제로 있을 때, 회선잡음의 효과를 평가하는 데는 오류를 범할 수 있다. 따라서, 실험계획단계에서 손실을 고려한 평가가 이루어지도록 신중한 주의가 필요하다.

송화 오피니언 시험(talking test)은 주로 에코와 측음에 의한 손실의 정도를 평가하기 위해 실시한다. 어떤 특성을 가진 전화 연결망에서 피험자는 1~2 개의 단문장을 송화하고, 품질척도 반응지에 '일상통화의 입장'에서 통화품을 평가한다.^{7,8)} 일반적으로 회화시험에서는 회화의 주제나 회화속도에 따라 에코의 정도가 변하기 때문에, 최근 디지털 전화기의 설계에서는 보다 엄격한 평정을 요구하는

송화시험을 채택하고 있다.⁹⁾

2. 실험실 평가와 현장 평가

실험실에서의 회화시험은 통화손실의 상대적인 주관효과를 나타내는데만 사용된다. 비록, 실험실 회화시험이 일반 전화사용자의 실제 현장 조건을 재현할 수 있다고 해도, 인위적으로 규정된 것일 뿐이다. 그러나 시험의 수행과 기술적 측면의 중대성을 고려하면, 실험실 회화시험은 적절하다고 볼 수 있다. CCITT 권고 P.11의 supplement No. 3과 4의 오피니언 모형은 이러한 목적으로 사용한다.

현장조건에서 통화품질의 실측치는, 일반 전화 사용자들에게 상용전화를 이용해서 통화하도록 한 후, 그들과의 면담으로 부터 얻을 수 있으며, 그때의 실험관리자 수행 방법이 CCITT 권고 P.77에 나타나 있다.¹⁰⁾ 그러나 다양한 목적을 위한 현장 관찰 방법은 평가하려는 전화 회선망에 대해 미소한 조정만을 할 수 있다는 단점을 가지고 있다.^{11,12)}

III. 시험 방법

2장에서는 오피니언 평가의 일반적인 측면에 대해 기술하였는데, 특히, 회화 시험과 청취 시험, 송화 시험 그리고 현장 시험과 실험실 시험간의 선택 지침을 제공하였다. 여기에서는 평가 계획에서 결과 분석까지, 오피니언 평가를 수행하는 절차에 관한 일반적 규칙과 주의사항을 기술한다. 또한 시험 목적에 따른 오피니언 평가척도를 4절에 나타낸다.

1. 청취 오피니언 평가

청취 오피니언 평가의 목적은, 임의로 조정된 음성 시료 샘플에 대해 통제된 환경(실험실)에서의 청취를 통하여, 통화품질에 대한 일반 전화 사용자들의 오피니언 체계를 확립하려는 것이다. 청취 오피니언 평가를 수행함에 있어서 실험자는 다음 사항들을 고려해야 한다.

가. 주관 평가 기준

통화품질 주관평가에서 얻은 결과는 각 시험조건에서 피험자가 사용한 기준에 의해 크게 영향을 받을 수 있다는 점을 인식해야 한다. 따라서 적절한 소의

니언 평가척도의 선별, 측정, 식도에의 범주(category)의 범주와 표현에 있어서 주의해야 한다.

일반적으로 사용하는 수화 평가척도는 전체적인 통화품질, 음량 선호성(loudness preference), 그리고 청취 노력(listening effort) 등이다. 일반적으로 5개의 범주를 사용하며, 범주의 범수가 다른 척도(예를 들면, 7개, 9개 등)도 있다. 다만, 보청, 청각이 범주간 관계성을 유지하면서 범주범수로 다른 언어로 번역할 때 상당한 주의가 필요하기 때문에, 오키니언 평가척도의 선별에 있어 국제적인 어떠한 모호적 지침을 제공한다는 것은 어렵지만, 국내에서는 1점에 나타난 척도를 사용한다.

나. 시험자의 선별과 수

시험자는 일반 전화 사용자의 대표적 모집단(population)에서 뽑힐당해야 하며, 통화적인 장벽을 가져야 한다. 또한, 적어도 6개월 이내에 유사한 주된 평가 시험에 참여한 적이 없어야 시험의 효과를 통제할 수 있다. 시험자수는 평가 시험의 정밀도와 함수관계를 갖기 때문에 신중히 선택하여야 하며, 일반적으로 30~50명 수준이 부닌하다.

다. 시험조건의 선별

오키니언 시험에서, 시험자의 판단 기준(reference of judgement)은 통화 시험내의 통화품질과 청취레벨의 변화에 의해 크게 영향을 받는데, 이러한 상승 효과(enhancement effect)라고 한다. 선택된 회선레벨이 지나치게 나쁜 상태를 많이 포함하지 않도록 하고, 최저치 이상에서부터 최저치보다 적어도 30 dB 미만의 범위까지의 청취레벨을 선정하고, 또한 각 수행 레벨내에서 적어도 한 집단의 음성자료 문장은 양호한 청취 레벨을 가진 고정조건(anchor condition)으로 제시되어 상승효과를 제거할 수 있다.¹²⁾ 시험 조건의 제시 순서는 일반적으로 순서효과(order effect)의 상승효과를 피하고자 하는 목적을 갖기 위하여 한다. 시험자의 통화품질 평가 기준을 초기화(initialization)하고 평가 절차에 익숙해지도록 하기 위하여, 유사 조건(similar condition)의 음향 조건을 평가자에게 보여 준다.

라. 음성자료의 준비

노래와 레벨에서 음향의 차이를 제공을 위해, 그리고 통제되지 않은 열화효과(degradation effect)의

영향을 방지 위해서 음성자료는 테이프에 미리 녹음된 것을 사용한다(가능하면, 디지털로 녹음된 것). 테이프 샘플의 수치를 최소화하기 위해 시험 준비기간 동안 고품질 디지털 테이프의 사용이 권고되었다. 최소한 2명과 이상 2명의 통화자들이 각 회선조건당 5초 정도 음향을 평가하여(약 2분간 분량), 약 20초 동안 그 자료를 통해 평가한다.

1) 시험 설계

시험절차에 대해 적절한 통계분석을 하기 위해, 시험설계는 회선조건, 시험자, 송화자, 그리고 제시순서의 유형을 고려해야 한다. 시험설계는 통상 그레코-라틴 방형설계(Graeco-Latin square design)나 라그레코-라틴 방형설계(hyper Graeco-Latin square design)를 기초로 하는데,¹³⁾ 횡렬은 청취자, 종렬은 조건제시 순서, 첫 열과 행은 회선 조건, 그리고 다음 열과 행은 송화자와 문장목록을 나타낸다. 설계의 각 cell은 하나의 수행을 나타내는 것으로 특정 송화자에 의해 녹음된 특정 문장목록을 특정 청취자에게, 특정 조건열(sequence of condition)에서 특정 회선조건을 통해 재생하는 것이다. 각 수행내에서, 청취레벨은 랜덤하게 미리 결정된 값에 따라 변화시키고, 시험자는 각 집단의 오키니언 평가척도의 한 범주에 표시한다.¹⁴⁾

바. 시험관리

오키니언 평가시험 관리에서는 다음 사항을 고려해야 한다.

1) 시험 자료물

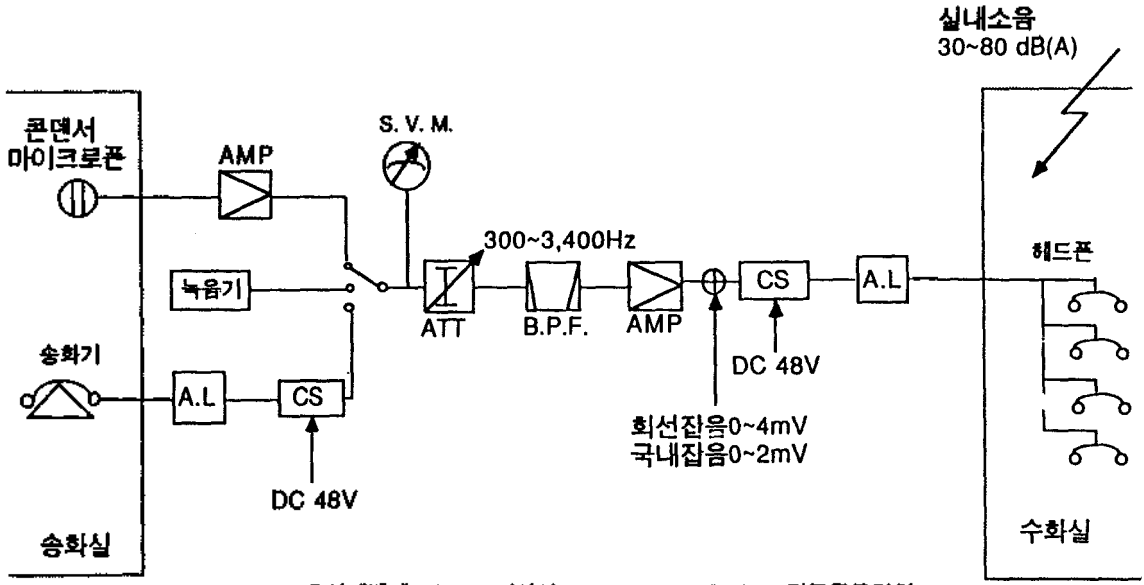
시험자에게 시험절차와 반응저장 방법 등 어떠한 의문점도 없도록 상세히 지시하고, 시험하고 있는 회선조건의 특성과 손실의 유형에 관해서 이야기해 주는 안된다.

2) 청취계(listening system)

콘드랫 전화기인, 콘소라티알 의식부하(dummy load)를 갖는 마이크로폰 캡슐로 교환 가능한 것을 사용해야 한다. 청취레벨은 통제된 전압적 신호를 발생하도록 고음역, 대 음향 척도보다 약 85dB SPL의 감호이다. 청취 레벨이다.

3) 청취 환경(listening environment)

정상적인 실내환경의 소음레벨은 35dB(A) 이하의 레벨이 적절하다. 실내소음 레벨을 가늠



S. V. M.: 음성레벨계, A. L.: 의사선로 (0~15dB), C. S.: 전류공급장치

그림 1. 청취 오피니언 시스템의 블록도

할 필요가 있으면 청취실에 여러개의 스피커 (전형적으로 2-4개)를 설치하여 확산음장(diffused sound field)을 만들어야 한다. 청취실의 크기에 따라 칸막이로 분리시켜 여러명 (4명 정도)의 피험자를 동시에 시험할 수도 있다.

4) 자료수집

피험자에게 각 조건에 대한 그들의 오피니언과 기타 관련된 질문을 반응지(response sheet) 기록하도록 한다. 학습이나 연합효과(learning or association effect)를 최소화하기 위해서는 한 반응지에 하나 이상의 반응을 기록하지 못하게 한다. 반응지에 기록하는 방식과는 달리 반응상자(responses box)의 버튼을 누르게 하는 자동 자료 기록법을 대안적으로 사용할 수 있다.

사. 시험 시스템

그림 1에 청취 시험 시스템의 블록도를 나타낸다.

2. 회화 오피니언 평가

청취 오피니언 평가와 비교해 보면, 회화 시험은 그 자체 특성상 실제 전화 통화에서 일어나는 자연

적 효과를 보다 정확히 재현한다. 송화나 청취 레벨에 어떠한 통제도 가하지 않고, 피험자에게 자유로운 회화를 유도한다. 다음에 회화 오피니언 평가의 일반적 절차와 주의사항들을 기술한다.

가. 통화 회선(telephone connections)

실험실내 전화회선은 피험자들이 실제 전화회선에서 경험하는 것과 같은 회로로 구성 되어야 한다. 축음이나 손실 등 회선 파라미터가 현실적이어야 하며, 특별하게 다른 종류의 핸드셋간 차이를 관찰하는 경우 이외에는 피험자가 정상시 사용하는 전화기와 외관, 갑축 등이 동일한 것을 사용한다.

나. 시험 지시문

시험 지시문은 명확하고 모호하지 않아야 하며, 회선의 구성, 회화 시험, 그리고 오피니언 실험 방법을 포함하여 시험의 모든 측면에 관한 사항을 다루어야 한다.

다. 실험설계

회화 오피니언 평가의 가장 적절한 설계 형태는, n명의 피험자들이 n개의 회선 조건 각각에 단 한번의 회화를 수행하는 n×n 그레드이다. 단 몇몇 선계를 사용한다. n이 8보다 작을 때, 실험의 정밀도가 떨어지기 때문에 n은 보통 8-15로 제한된다.

일반적으로 그레코·라틴 방형설계에서, 통상 횡렬은 피험자쌍, 종렬은 실험실시 순서, 첫 알파벳은 회선 자체의 특성 뿐만 아니라 실내 소음 레벨과 다른 처치요인에 따라 구별되는 회선 조건을 그리고 두번째 알파벳은 회화의 주제로 사용되는 시험시료를 나타낸다. n=10, 14 또는 15 일때, 더 이상의 독립적 요인(orthogonal factor)이 결합될 수 없으며, n=8, 9, 11 또는 13일 때, (n-3)개의 부가적 알파벳을 사용하여 과그레코·라틴 방형설계의 구성이 가능하다. 이것은 각 회화에 대해서 또 다른 독립적 요인(예를 들면, 탄소 마이크로폰의 선택이나 누화 녹음(crosstalk recording)의 선택)을 다루는데 사용한다. 일반적으로 사용되고 있는 n은 12이지만 현재 권고된 n값은 13이다.

다. 회선조건의 선택

몇개의 회선조건 간에 특별히 정밀한 비교가 필요한 경우, 그 회선조건들은 동일한 실험내에 속해야 한다. 또한 모든 실험에서 피험자들에게 전체 수행레벨 범위를 경험하도록 하기 위해서, 시험조건 범위는 충분히 넓어야 한다. 즉, 적어도 하나의 매우 좋은 회선 조건, 거의 평균에 가까운 회선조건, 매우 나쁜 시험조건이 있어야 한다. 이에 대한 신뢰성을 보장하기 위해서, 선택한 회선 조건 세트에 대해 실험 실시전에 간단한 예비시험을 수행하여 선택한 회선조건 범위가 타당한지를 검토하는 것도 인고되어 있다. 신뢰성을 보장할 수 없을 때, 피험자의 오피니언 평가치는 상승 효과와 같은 왜곡된 형태를 보일 수 있다. 실험자가 직접 관심을 갖고 있지 않은 여분의 회선조건(extra conditions)이 가끔 부가되는데, 이것은 실험 설계상 n을 9, 11 또는 13으로 만들기 위해, 또는 수행 범위의 보다 효과적인 균형을 위해 필요하다. 시험 초기에 각 피험자쌍에게 시험진

행에 익숙해 지도록 하기 위해 약 3개의 의사조건(dummy condition)을 제시하여 초기 통화품질의 기준을 형성하도록 유도한다.

마. 피험자의 선택

피험자의 선택은 창취 오피니언 평가의 경우와 동일하다. 피험자의 수는 12×12 라틴 방형설계에 준거한 실험설계를 광범위하게 사용하는데, 24명의 피험자(12 피험자쌍)와 의사조건을 제외한 12시험조건이 필요하다. 실험설계에서 남/여 피험자수의 조절을 요구하지 않는 한, 어떠한 남/여 피험자의 선발절차도 필요하지 않다. 피험자는 실험 설계에 따라 임의적으로 짝짓고, 시험실시 기간동안 계속 그 쌍을 유지한다.

바. 회화 환경(conversation environment)

창취 오피니언 평가에서와 같이 주변조건이 통제되고, 음향적으로 분리된 두개의 방음실이 필요하다. 실내소음은 일반적으로 통제된 레벨(보통 50dB(A))의 Hoth잡음을 사용하며, 이 소음레벨은 sound level meter(예를 들면 B&K Precision Sound Level Meter type 2206)로 측정한다. 만일, 필요에 의해 실내소음을 변화시킬 경우에는 피험자를 방음실 밖으로 내보낸 뒤 가변하는 것이 바람직하다.

사. 회화시험 내용

피험자의 회화가 어떠한 의미부족 갖도록 하고, 시험 회선의 통화품질에 대한 그들의 평가에 모든 기회를 제공한다 것은 상당한 노력이 필요하다. 어떤 가치 순서에 따라 그림을 재배치하는 시험이 이러한 목적에 적절하며, 피험자에게도 상당한 흥미를 준다(예를 들면, 그림 2의 그림 시료). 다양한 주제를 담고 있는 표준 업서크기의 삽화 견본으로 카드를 만들고, 이 카드 뒷면에 개별번호를 매겨, 각각 6개씩 모아 하나의 세트를 만든다. 시험회선을 통해 대화하

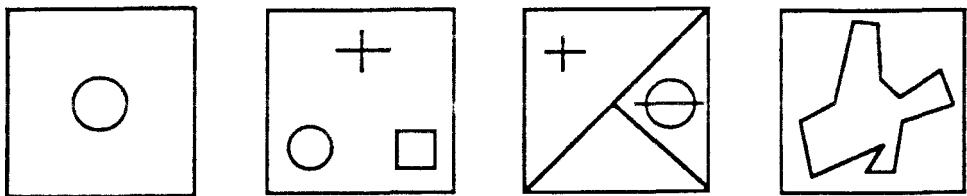


그림 2. 회화시험에 사용하는 그림 세트 예

하기 전에, 피험자들의 개인적인 선호에 따라 카드를 배열하도록 지시한다. 시험회선으로 접속이 이루어지면 피험자들은 자신들의 선호 순서에 관해 토의하고, 회화가 끝나면 이 결과를 반응지에 기록한다. 각 회화의 지속시간은 피험자들에게 일임한다.

아. $n \times n$ 실험 준비

실험자는 선택된 피험자에게 참가시간과 약속장소 등을 전화로 예약해 두고, 실험에서 피험자에게 요구되는 몇가지 세부사항을 설명하는 편지를 보낸다. 또한 실험자는 실험설계를 기초로, 각 피험자쌍에게 어떤 제시순서를 적용하는가, 어떤 그림세트를 사용하는가, 각 피험자쌍에서 누가 통화를 시작할 것인가 등 기타 필요한 세부사항에 관한 시간표를 작성한다. 편지나 실험실시 중, 회선조건과 특성에 대해서는 어떠한 이야기도하면 안된다는 사실을 항상 주의해야 한다. 회화시험에 대해서는 문헌¹⁰⁾에 자세히 기술되어 있다.

자. 실험절차

피험자가 실험실에 처음 방문했을 때, 실험자는 편지를 완전히 이해 했는지 물어보고, 모호한 점이 있다면 명확한 설명을 해 주어야 한다. 실험에 사용하는 방음실과 장치 등을 보여주고, 이번 방문에서 피험자가 해야 할 통화의 횟수를 알려주고 반응지를 나누어 준다.

회화시험에 들어가기 전에, 각 피험자는 책상위에 있는 상자에서 특정 그림세트를 선정하고 자신의 선호에 따라 순서를 배열한다. 그 배열 결과를 오피니언 반응지의 해당 부분에 기입한다. 두 피험자 모두 반응지에 기입을 끝내면, 실험자는 둘중 한명에게 전화개시의 신호를 보낸다. 전화 회선에 대한 의견을 교환해서는 안된다는 점을 제외하고는, 대화는 완전히 자유롭게 진행시킨다. 피험자들이 그림에 대해 일정한 선호순서를 기입하고, 또한 통화품질에 대한 그들의 오피니언을 묻는 항목에 기입한다. 다음에 회선을 통해서 피험자 자신이나 파트너가 듣거나 말할 때 어려움이 있었는가 하는 난이도 문항에 '예 / 아니오' 반응을 하도록 하고, 만일 답이 '예' 일 경우, 피험자에게 이유를 간단히 설명하도록 질문한다.

첫 그림세'에 대한 실험이 끝나면, 실험자는 피험

자에게 반응지를 봉투에 넣도록 지시하고, 다른 그림 세트를 섞게 하거나, 경우에 따라 방음실에서 나와 기다리도록 한다.

피험자간 대화와 피험자-실험자간 대화는 전부 녹음테이프에 저장한다. 회화시험에서는 피험자의 음성레벨에 어떠한 통제도 하지 않으며, 송화 음성레벨과 회선 파라미터간의 상관 관계는 상당히 중요하기 때문에 회화 시험 동안에 녹음된 음성레벨에 대한 자료는 차후에 유용한 목적에 사용하기 위해 분석할 수 있다. 이러한 사실은 피험자에게 절대 비밀로 해야 한다.

차. 회화 시험 시스템

실험 목적에 따라 2선식 또는 4선식 시스템을 선택하는데, 그림 3에는 2선식 시스템을, 그림 4에는 4선식 시스템을 나타낸다.

3. 송화 오피니언 평가

송화 오피니언 평가는 실험실에서 실험자가 임의로 가변한 전화망을 사용하여, 에코와 측음 등에 의해 발생하는 회선 손실의 정도에 따라 일반 전화 사용자들의 오피니언 체계를 확립하는데 목적이 있다. 청취나 회화 오피니언 평가의 시험방식과 중복되는 부분은 제외하고, 송화 오피니언 평가를 수행함에 있어 실험자는 다음 사항들을 고려하여 실시하여야 한다.

가. 피험자의 선발과 수

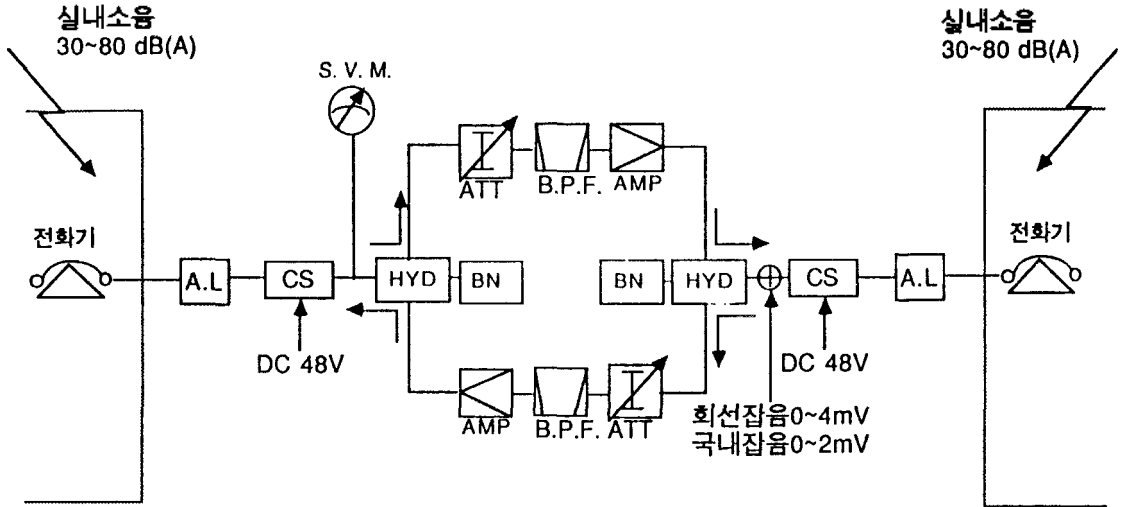
피험자의 선정은 청취 오피니언 평가에서의 방법과 동일하다. 통상적으로 실험에서는 핸드셋 사용방식과 발생레벨에 대한 특별한 지시를 하지는 않지만, 정상인과 비교해 극단적인 음색을 가진 자(예를 들어, 지나치게 쉼 목소리나 코막힌 소리 등)를 피험자로 사용하는 것을 가급적 피하는 것이 좋다.

나. 실험설계

실험설계는 회선조건, 송화자, 그리고 실험 제시물의 유형을 고려하여, 통상 라틴 방형 설계(Latin square design)를 사용한다¹¹⁾.

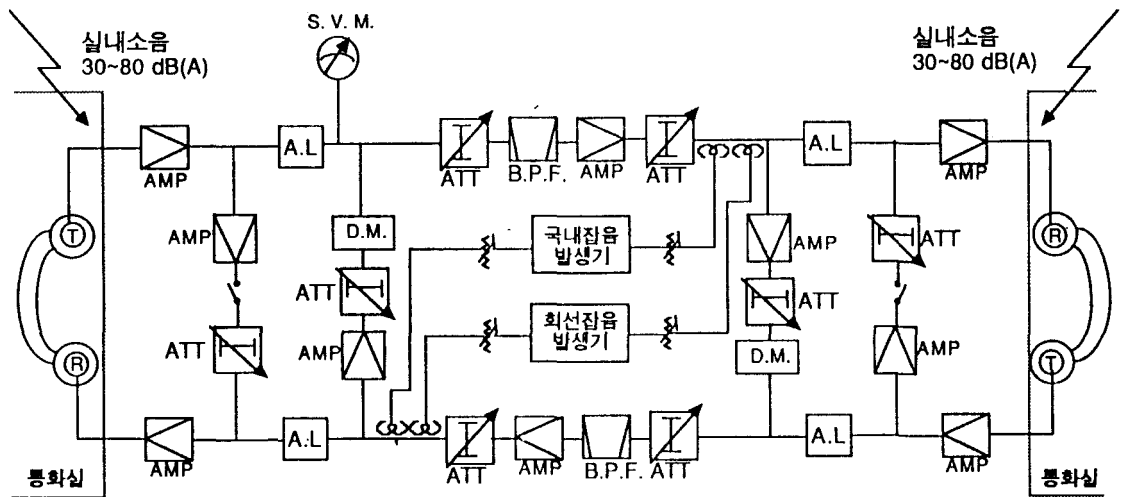
다. 송화 환경

송화 시험에서의 실내소음은 일반적으로 15dB



S. V. M. : 음성레벨계, A. L. : 의사선로 (0~15dB), C.S. : 전류공급장치, HYD : 하이브리드, BN : 평형 결선망
AMP : 증폭기, B. P. F. : 대역통과 필터, ATT : 감쇠기

그림 3. 2선식 회화 오퍼니언 시스템의 블록도



S. V. M. : 음성레벨계, A. L. : 의사선로 (0~15dB), C.S. : 전류공급장치, HYD : 하이브리드, BN : 평형 결선망
AMP : 증폭기, B. P. F. : 대역통과 필터, ATT : 감쇠기, D. M. : 지연기

그림 4. 4선식 회화 오퍼니언 시스템의 블록도

A) 정도의 Hoch 잡음의 작용하여, 그 음레벨의 일정
- 관찰하기 위해서는 30~80dB(A)까지 가변한다
리, 중화 지교
중화 시험에 사용되는 지교는 한 조건일, "작위함

수없이 부하와 다른으로 3000Hz와 같은 1~2개
다른 것(약 5중 정도)을 기용한다.
의, 중화 시험 시스템
그림 5에 중화 시험 시스템을 나타낸다.

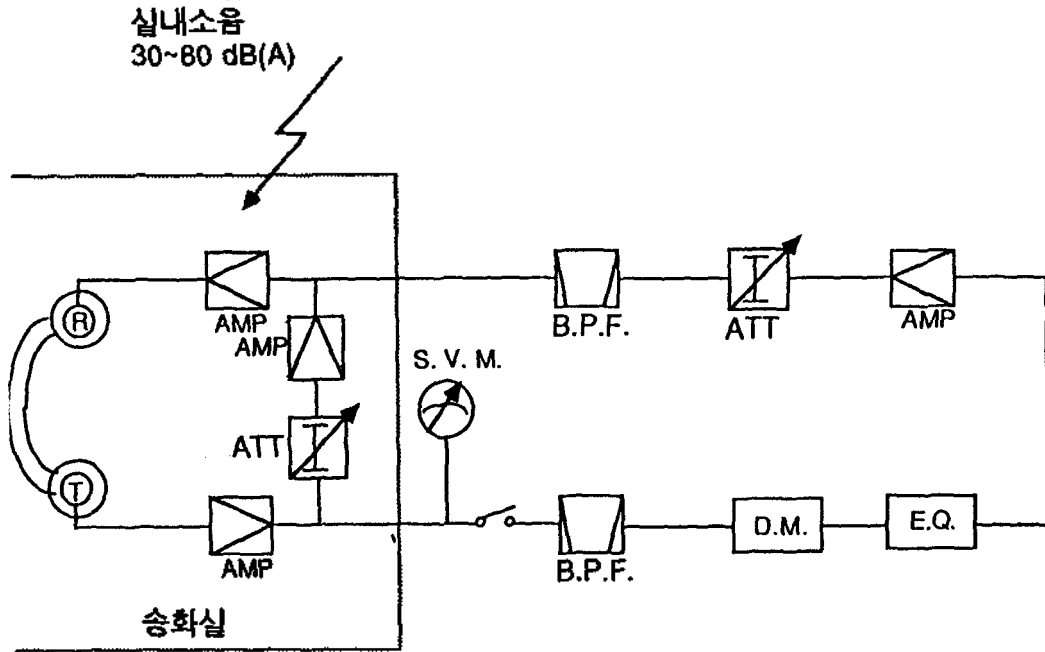


그림 5. 송화 오피니언 시스템의 블록도

4. 오피니언 평가척도

가. 전체적인 품질 오피니언 척도

- A. 아주 좋다.
- B. 좋다.
- C. 보통이다.
- D. 나쁘다.
- E. 아주 나쁘다.

나. 오피니언 척도 (No.1)

- G. 아주 좋은 회선: 어떠한 노력도 불필요
- F. 좋은 회선: 회화에 약간의 노력 필요
- P. 나쁜 회선: 회화는 가능하나, 낯선 단어는 청취 어려움
- B. 아주 나쁜 회선: 거의 회화가 불가능

다. 라우드니스 선호성 척도(No.4)

- A. 아주 크다
- B. 크다
- C. 적당하다
- D. 작다
- E. 아주 작다

라. 라우드니스 선호성 척도(No.4A)

- A. 선호하는 정도보다 아주 크다
- B. 선호하는 정도보다 크다
- C. 선호하는 정도이다.
- D. 선호하는 정도보다 작다
- E. 선호하는 정도보다 아주 작다

마. 라우드니스 수용성 척도(No.4B)

- 4. 너무 크다
- 3. 수용하기에 크다
- 2. 적당하다
- 1. 수용하기에 작다
- 0. 너무 작다

바. 오피니언 척도 (No.5)

- A. 들리지 않는다: 잡음을 감지할 수 없음
- B. 겨우 들린다: 주의깊게 청취하면 감지 가능
- C. 약간 들린다: 그러나 방해되지 않음
- D. 보통이다: 잡음이 약간 방해
- E. 조금 크다: 잡음이 방해

음성품질 주관 평가법

- F. 크다: 잡음이 상당히 방해하지만, 통화는 계속 가능
- G. 참기 어렵다: 잡음이 너무 커서 통화 불가능
- 사. 양자-반응 검지가능성 시험(Quantal-response detectability tests) (No.6)
- A. 불쾌하다
- B. 지각할 수 있다
- C. 지각할 수 없다
- 아. 양자-반응 검지가능성 시험(No.6A)
- A. 불쾌하다.
- B. 검지할 수 있다
- C. 검지할 수 없다.
- 자. 청취 노력 척도(No.7)
- A. 아주 편한 느낌으로 통화가능: 어떤 노력도 불필요
- B. 상대방의 이야기에 주의가 필요 하지만 노력은 불필요
- C. 적당한 노력 필요
- D. 상당한 노력 필요(통화는 가능하지만, 익숙하지 않은 단어는 알아 듣기 어려움)
- E. 상당히 노력하지 않으면 의미 이해 불가능
- 차. 회화 노력 척도(No.8A)
- A. 아주 편한 느낌으로 통화가능: 어떤 노력도 불필요
- B. 상대방의 이야기에 주의가 필요 하지만 노력은 불필요
- C. 적당한 노력 필요
- D. 상당한 노력 필요(통화는 가능하지만, 익숙하지 않은 단어는 알아 듣기 어려움)
- E. 상당히 노력하지 않으면 회화 불능

IV. 결과 분석

오피니언 점수가 주어진 후에, 주관성결과를 분석한다. 타당할 결론을 내리는데 적절한 형태로 그것들을

제시한다. 주관적 오피니언 평가는 그 자체 특성이 매우 가변적이기 때문에, 시험 결과에 대한 분석은 실험설계에 따라 적절한 통계방식을 적용해야 한다. 오피니언 평가 결과는 일반적으로 평균 오피니언 점수(mean opinion score: MOS)로 분석하며, 때로는 각 오피니언 범주의 누적 선호도(cumulative preference, Σ)를 이용한다.

예를 들어, 그림 6은 5점 오피니언 척도에서 다섯 개의 범주 A, B, C, D, E에 각각 5, 4, 3, 2, 1(또는 4, 3 2, 1, 0)의 가중치(weight)를 부여한 경우를 나타낸다. 평균 오피니언 점수 Y는 각 범주에 가중치를 주어 다음의 식으로 계산한다.

$$Y = \sum_{i=1}^n p_i x_i \quad (1)$$

여기서, P_i : i 번째 범주를 선택한 확률

x_i : i 번째 범주에 주어질 가중치

$x_i=0\sim 4$ (또는 $1\sim 5$)

n : 범주의 수 (여기에서는 $n=5$)

표준편차와 신뢰구간(confidence interval)은 통상적인 공식으로 계산할 수 있다. 그러나, MOS는 그림 6과 같이, 순서에 대한 함수로 표현하여 '평균' 주관적 오피니언으로 나타내지만, 오피니언 범주에 할당된 명명(label)이 반드시 등거리(equidistance)를 나타내는 것은 아니기 때문에, 간단히 수치적 평균을 하는 것에 타당성의 문제를 제기할 수 있다. 그럼에도 불구하고, MOS는 주관 시험의 결과를 해석하는데 실질적으로 매우 유용하다고 판명된 일반적 기법이다⁶⁾.

등거리 명명의 문제점을 해결하기 위한 다른 분석 방법은 누적 점수(cumulative score)에 의한 계산법이다. 누적 선호도를 노출하기 위해 각 실험조건에 대해, 주관 반응을 각 범주에 대한 오피니언의 %점수로 하고, 가로축을 누적확률 척도로, 세로축을 시험점수로 하여 표시적으로 나타낸다(그림 7). 가끔 그림 8과 같이 Good나 Excellent 범주, Poor와 Bad 범주로 전체 반응을 통합하기도 한다. 일반적으로, % GOS(% Good or Better)와 % POW(% Poor or Worse)로 나타내며, 매우 좋거나 매우 나쁜 회선

조건에 대해서는 MOS보다 다소 엄밀한 지수로써 사용된다.

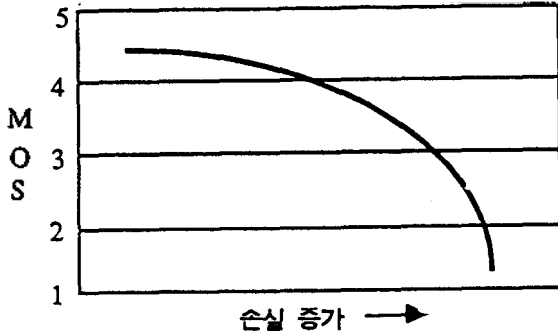


그림 6. 가중치를 부여한 오피니언 척도

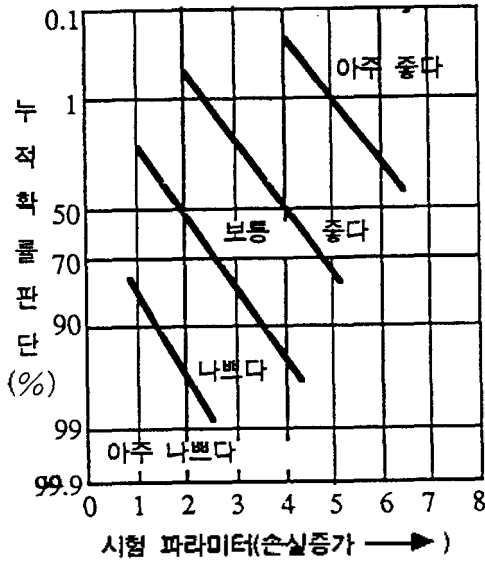


그림 7. 주파수 대역도

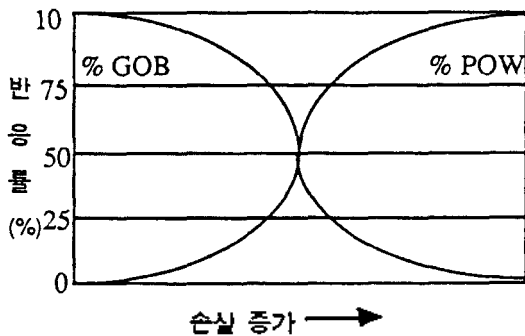


그림 8. 통행된 전체 비율도

V. 응용 사례

1. 오피니언 평가와 통화당량과의 관계

임의의 전화 전송계에서 통화당량(RE)을 파라미터로 하여 전송품질을 나타내고자 하는 경우에 평균 오피니언 점수와 RE와의 관계를 유도할 수 있다. 그림 9는 송화계 가입자선 0.4mm ϕ -3dB, 수화계 가입자선 0.4mm ϕ -6dB, 실내 소음레벨을 60phon으로 고정하고 회선잡음을 파라미터로 한 경우의 RE와 평균 오피니언 점수와의 관계를 나타내고 있다.

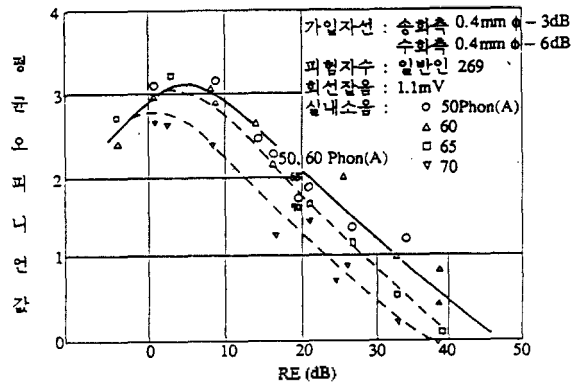


그림 9. 평균 오피니언 점수와 RE와의 관계

2. 오피니언 평가와 명료도 증가 감쇠량과의 관계

그림 9와 같은 조건에서, 평균 오피니언 점수와 명료도 증가 감쇠량(AEN)과의 관계를 그림 10에 나타낸다.

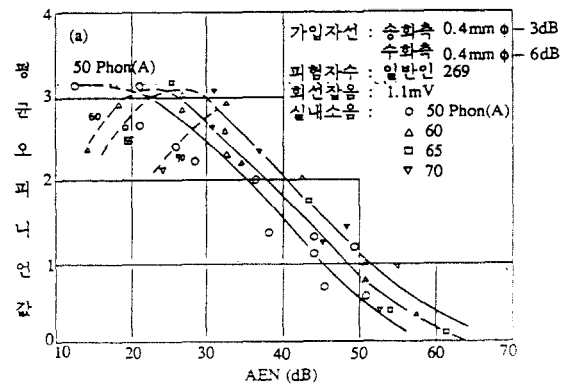


그림 10. 평균 오피니언 점수와 AEN과의 관계

3. 오피니언 평가와 측음과의 관계

RE의 변화에 따라, 수화계 가입자선이 0.4mm ϕ -6dB 및 0.5mm ϕ -9dB 일 때의 평균 오피니언 점수를 그림 11에 나타낸다. 그림 11의 경우, 통화 시험원에서만 측음효과가 크게 나타낸다.

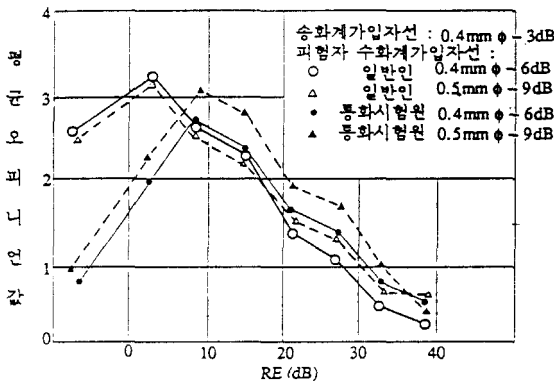


그림 11. 평균 오피니언 점수와 측음과의 관계

4. 기타 응용분야

앞에서 제시한 분야외에, 오피니언 평가는 회선의 선반직 음질, 라우트리스 선호성, 라우트리스 수용성 (loudness acceptability), 소음, 누화 등의 검출 (detection), 청취노력 (listening efforts), 회화노력 (conversation efforts) 등의 인간에 관련된 하계 사용 의미, 디지털 전화기를 비롯한 ISDN 관련 통신기기의 통화품질 기준설정에도 응용할 수 있다.

VI. 결 론

지금까지 CCITT 보고 자료를 중심으로 음성품질 주관평가법의 표준화를 위해, 주관평가법 중 가장 많이 사용되고 있는 오피니언 평가법에 대해 살펴보았다. ETRI에서는 앞으로 음성품질 주관평가법 연구에서 기술한 방식을 토대로 실험할 예정이며, 음성품질 주관평가법 연구에 많은 관심을 바란다.

회 통화질의 저원자에 대한 통화품질 표준 수준을 전화로 평가하고, 음대 수화시험원 직감의 진위 여부를 인정하기 위해서는 우리의 실정에 맞는 음성품질

주관 평가법의 표준화가 절실히 요구된다. ETRI에서는 통신에 있어서 음성품질 주관 평가법의 표준화¹⁾와 함께 음성품질 표준 평가 시스템을 개발 중에 있으며, 이에 대한 결과는 시스템 개발이 완료되는 데로 보고할 예정이다.

참 고 문 헌

1. 日本音響學會明瞭度委員會: 明瞭度試驗法の規準 (1957).
2. 飯田: 明瞭度 테스트における問題点. 音響技術 7, July, 1974.
3. CCITT Recommendation P.11 Effect of transmission impairments, Yellowbook, Vol.V, Geneva 1980.
4. CCITT Recommendation P.74 Methods for subjective determination of transmission quality, Vol.V.
5. J.L. Sullivan, "Is Transmission Satisfactory?," Bell Lab. Record, pp. 91-98, March, 1974.
6. CCITT Recommendation G.102 Transmission performance objectives and Recommendations, Vol. III, Fascicle III.1.
7. 井谷, 入井, "最適な側音特性の検討." 日本音響學會音響研究資料, S81-33 (1981).
8. CCITT Contribution 116 Desirable Sidetone Masking; Rating for Digital Telephone Sets, Study Group VII COM VII-No 9 and 10, Sept, 1986.
9. 井谷, 入井, "送話時 手話時の通話品質に及ぼす側音の影響," 日本音響學會 S79-36(1979)
10. CCITT Recommendation P.77 Methods for evaluation of service from the standpoint of speech transmission quality, Yellow Book, Vol. V, Geneva 1980.
11. HELDER(G.K.): Customer evaluation of telephone circuits with delay. BSTJ, September 1966, pp. 11-570, 1191.
12. POST(J.A.), SILVERTHORN(R.D.): Results of a subjective comparison of echo control devices in terrestrial and satellite trunks. Part C of NTC Recommendation, Vol.3, November 1979.
13. CCITT Recommendation P.74 Methods for subjective determination of transmission quality, Vol. V.
14. 김정환, 강성훈, "통신에 있어서 음성품질 주관 평가법", 한국통신학회지, TMS96-1230-01 (1996).
15. RICHARDS(D.L.): Telecommunications by speech. Butterworth & Co., 1973.
16. CCITT Recommendation P.12 ANNEX 2, Yellow Book, Vol. V, ITU Geneva 1981
17. 辻藤 "電話傳送品質の客観的評価についての検

討, 研究實用化報告 21, 2, 1972.

筆者紹介

▲김정환



1964년11월15일생
1987년2월: 고려대학교 심리
학과 졸업(문학사)
1989년2월: 고려대학교 대학
원 심리학과졸업
(문학석사)
1989년2월~현재: 고려대학교
행동과학연구소
연구원

1989년11월~현재: 한국전자통신연구소 신호처리연
구실 위촉연구원

▲김민택



1955년8월26일생
1979년2월: 아주대학교 전자
공학과 졸업 (공
학사)
1984년2월: 연세대학교 대학
원 전자공학과졸
업(공학석사)

1984년7월~1985년7월: 아주대학교 전자공학과전
임조교

1985년7월~현재: 한국전자통신연구소 신호처리연
구실 선임연구원

▲강성훈 한국음향학회지 제9권1호 참조.