

청각 보철을 위한 자극 패턴 추출에 관한 연구

박 상희, 윤 태성, 이재혁, 백 승화*
 연세대학교 전기공학과 영지대학교 전기공학과

A Study on the Extraction of the Excitation Pattern
 for Auditory Prosthesis

Sang Hui Park, Tae Sung Yoon, Jae Hyuk Lee, Seung Hwa Beack
 Dept. of Electrical Eng., Dept. of Electrical Eng.,
 Yonsei Univ. Myongji Univ.

In this study, the excitation pattern, which can be sensed by a man having hearing loss due to the damage of inner ear, is extracted, and the procedure of the auditory speech signal processing is simulated with the computer.

Therefore, the excitation pattern is extracted by the neural tuning model satisfying the physiological characteristic of the inner ear and by the information extracted from speech signal. The firing pattern is also extracted by inputting this excitation pattern to the auditory neural model. With this extracted firing pattern, the possibility that the patient can sense the speech signal is studied by the computer simulation.

1. 서 론

음성에 대한 지각이나 인식을 일으킬 수 있는 자극에 대한 연구는 1970년대부터 진행되고 있으나, 지각을 일으키는 확실한 자극패턴은 얻지 못하였으며, 보철신호는 내이에 주입된 전극을 통하여 음성신호와 유사한 패턴으로 기저막 상의 신경을 자극하는 것이 일반적이다.

본 연구에서는 내이의 손상에 의하여 청각장애가 있는 사람에게 음성에 대한 지각을 일으킬 수 있는 적당한 자극패턴을 신경동조 특성과 음성정보로부터 추출하며, 내이에서의 음성정보 처리과정을 시 레이선하였다.

그러므로, 본 연구에서는 음성 중에서 한국어 모음에 대한 특성을 추출하였으며, 음성에서 추출된 정보들 내이의 정보 전달경로를 따라서 신경동조 모델에 적용하여 청각신경 모델에 입력되는 자극패턴으로 만들고, 이 자극을 청각신경 모델에 입력하여 나타나는 음성신호에 대한 응답을 고찰하고자 한다.

신경동조 특성은 Rhode의 측정치와 일치하는 한 쌍의 미분방정식의 해를 구하여, 기저막 압력과 변위에 대하여 급격히 동조하는 모델로부터 구한 것이며, 청각신경 응답은 신경의 세포막에서 이루어지는 이온의 변화에 따라 활동전위가 발생하는 Hodgkin-Huxley 모델로부터 얻을 수 있다.

본 연구의 대상은 청각기관 중 헤어셀 이하의 모든 청각기관이 손상되고 오직 청각신경만이 살아있는 사람으로 하였으며, 자극패턴 추출에 대한 검증은 기존의 실험자료와 비교하였다.

2. 내이의 특성과 청각 신경의 모델

중이의 진동에 의해 등골(stapes)이 피스톤 운동을 일으키면 코클리어 내부 액체의 압력차에 의해 기저막의 변위가 발생된다. 이때, 기저막의 변위가 최대인 지점은 주파수와 관계가 있다. 즉, 고주파의 소리에 대해서는 스테이프스 근처에서 최대변위가 발생하며 주파수가 낮아질수록 최대변위의 발생장소는 헬리코트

레마 쪽으로 이동한다. 따라서 기저막은 주파수 해석기 또는 대역통과 필터의 기능을 갖는다고도 볼 수 있다.

헤어셀의 주된 기능은 유입신호의 기계적인 압을 신경활동으로 변환시키는 것이다. 기저막의 운동이 일어나면 헤어셀의 섬모가 구부러지거나 당겨지는 등의 변형이 일어나, 그에 연결된 신경섬유에 신경 임펄스가 발생하게 된다.

대개의 청각신경 섬유의 경우, 여러 주파수를 갖는 소리의 자극에 대하여 반응을 보이므로, 한 신경섬유에 여러 주파수를 갖는 소리의 자극을 그 크기를 달리 해서 가해보면 각 주파수에 대한 자극 문턱치를 알 수 있다. 그림 2-1은 두 신경섬유에 있어서 주파수와 자극 문턱치의 관계를 나타낸 곡선이다.

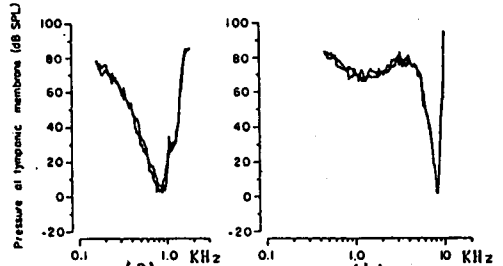


그림 2-1 두 신경 섬유의 동조 곡선

그림 2-1의 곡선을 그 신경의 신경동조 곡선이라 부른다. 한 신경이 SR(spontaneous rate) 이상으로 발화하게 만드는 최소 크기를 갖는 주파수를 그 신경의 CF(characteristic frequency) 혹은 BF(best frequency)라 한다. 신경동조 곡선의 최소점이 바로 CF에 해당한다. 그림 2-1(a)의 신경섬유의 CF는 대략 0.9KHz 이고, 그림 2-1(b)의 신경섬유의 CF는 9KHz 정도인 것을 알 수 있다.

중추신경계에서의 음성신호에 대한 처리는 청각신경의 발화패턴에 의하여 결정된다. 발화패턴은 특정 장소에 대한 발화율과 시간의 변화에 따른 발화율로 나타낸다.

Young과 Sachs에 의해 측정된 모음 /E/에 대한 청각신경에서의 발화패턴은 그림 2-2와 같다. 그림 2-2에서 음성신호의 포먼트 주파수 근처에서 발화율이 크게 나타나는 것을 보여주고 있다.

청각신경 응답은 신경의 세포막에서 이루어지는 이온의 변화에 따라 활동전위가 발생하는 Hodgkin-Huxley 모델로부터 얻을 수 있다.

그림 2-3은 30 μ A와 140 μ A의 자극에 대한 신경응답이다. 그림 2-3(b)의 발화율은 약 99 spikes/sec이다. 여기서, 그림 2-3(b)는 그림 2-3(a)보다 활동전위의 최대값이 많이 감소한 것을 볼 수 있다. 자극전류의 강도가 커질수록 발화율은 증가하나 활동전위의 최대값은 감소하는 현상을 보여 주고 있다.

그림 2-4는 자극전류의 세기에 따른 평균 발화율의

관계를 나타낸다. 반복성 발화율 위한 자극전류의 문턱치(약 6.4 μ A)로서 이보다 작은 전류를 가했을 때에는 반복적인 활동전위를 볼 수 없다. 6.4 μ A 이상의 자극에 대해서는 전류의 세기가 증가함에 따라 평균 발화율도 증가함을 알 수 있다. 그러나, 자극의 세기가 어딘 일정한 한도(여기서는 140 μ A) 이상을 넘으면 그림 2-3(b)와 같이 반복되는 활동전위의 크기가 현저히 감소한다(이 상한선 이상을 포화영역이라 한다.) 평균 발화율에 고려하지 않았다.

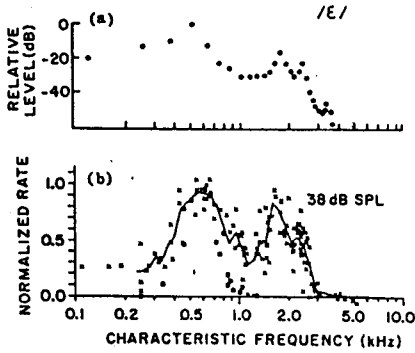


그림 2-2 모음 /E/에 대한 발화 패턴

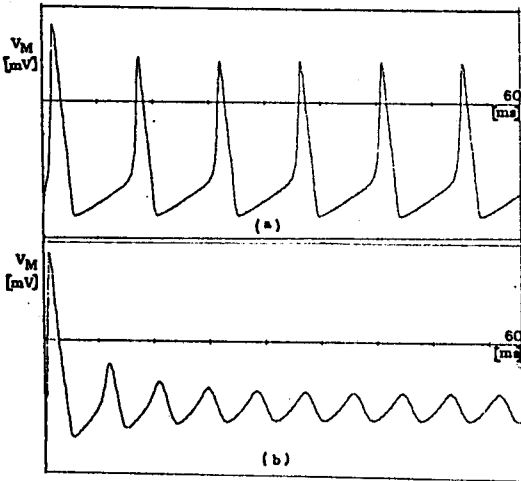


그림 2-3 신경 응답

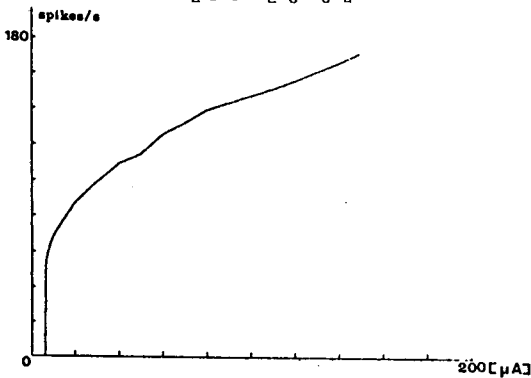


그림 2-4 활동 전위의 발화율

3. 자극패턴의 추출

일정한 크기의 주파수가 내이에 입력되었을 때 기저막 상에서의 자극형태는 신경동조 모델의 곡선과 같이

특성주파수 위치에서 가장 큰 자극이 가해지게 된다. 따라서 음성신호가 내이에 입력되면 그 신호의 주파수 성분(포먼트 주파수)은 신경동조 곡선으로부터 자극패턴을 결정할 수 있다. 음성신호와 신경동조 곡선으로부터의 자극패턴 추출에 대한 흐름도는 그림 3-1과 같다.

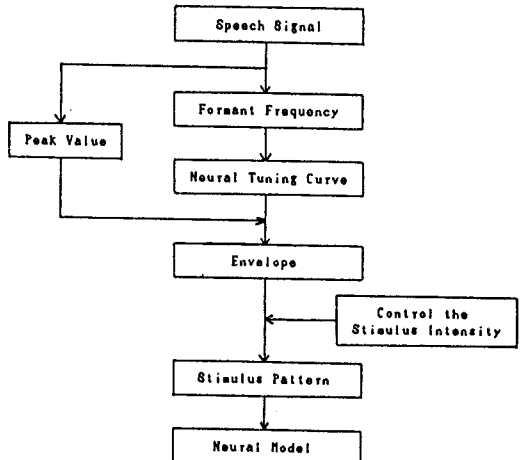


그림 3-1 자극패턴 추출의 흐름도

음성신호로부터 추출된 포먼트 주파수는 신경동조 모델에 입력되어 기저막 상의 특정위치와 그 부근에서 동조된다. 또한 각 포먼트 주파수는 그의 진폭특성이 모두 다르므로 규준화(normalize)되어 각 동조 곡선에 곱하여진다. 이렇게 하여 얻어진 포락선(envelope)은 내이에서 청각신경에 가해주는 자극의 크기로 조절된다. 본 연구에서는 전극에 의하여 기저막 상의 모든 신경에 자극을 주입할 수 있는 이상적인 보철 모델로 가정하였다.

본 연구에서는 청각신경에 입력되는 자극신호 패턴을 만들기 위하여 한국어 모음을 대상으로 선형 예측계수를 이용하여 포먼트를 추출하고자 한다. 음성 파형, 특히 경성적인 모음의 경우 그 스펙트럼 상에서 특징의 주파수 영역에 에너지가 집중해서 나타나는 것을 알 수 있는데 이것을 포먼트(formant)라 하고, 그 영역의 중앙치 또는 최대값을 포먼트 주파수라고 한다.

코클리어 모델은 일반적으로 이상적인 액체로 가득 찬 각형 모델로 가정할 수 있으며, 신경동조 모델은 기저막에 대한 Rhoads의 측정치와 일치하는 한 쌍의 미분방정식의 해를 구하여, 기저막 압력과 변위에 대해 급격히 동조하는 모델로 구현하였으며, 그림 3-2는 입력 스테이프스 속도의 주파수를 저주파수에서 고주파수 단계적으로 변화시켰을 때, 기저막의 각 지점에서 신경동조 응답과 기저막 속도 응답의 진폭특성을 나타낸다. 스테이프스에 가까운 곳일수록 고주파수의 공진이 일어나고, 헬리코그래프에 가까운 곳일수록 저주파수의 공진이 일어나는 이동파 특성을 나타내며, 신경동조 응답은 기저막 응답에 비해 더 급격하게 동조함을 보여준다.

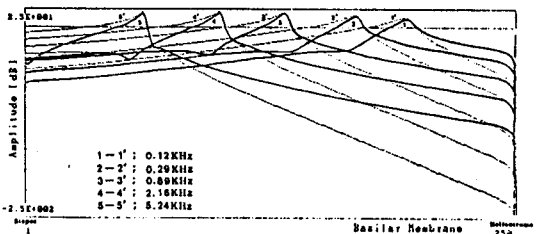


그림 3-2 주파수 변화에 따른 신경동조 응답(1-5) 및 기저막 속도 응답(1'-5')

4. 실험 및 결과 고찰

본 연구에 대한 전체 실험의 시스템 구성도는 그림 4-1과 같다.

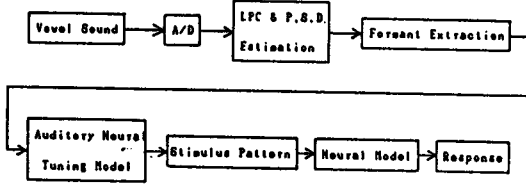


그림 4-1 전체 실험 시스템의 구성도

실험에 대한 검증은 Young과 Sachs가 실험한 모음의 주파수 특성에 대한 데이터로부터 본 실험의 모델-신경동조 모델, 청각신경 모델-에 적용하여 신경에서 실제로 측정된 신경 발화패턴과 모델로부터 얻어진 결과를 비교하였다.

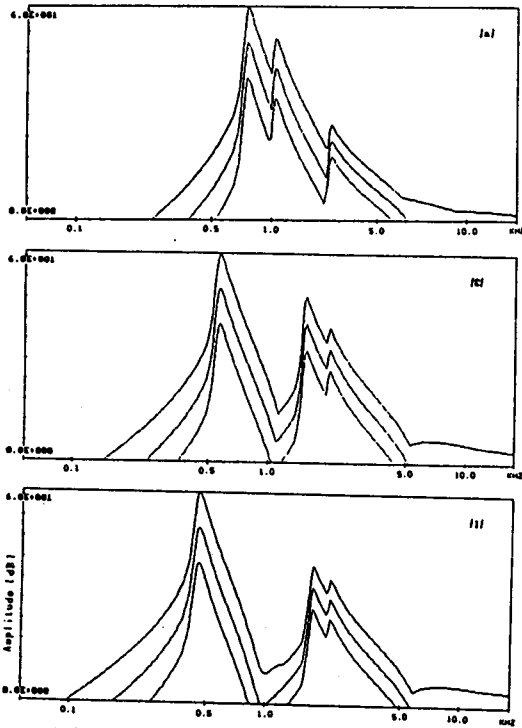


그림 4-2 40,50,60dB에 대한 신경 자극패턴

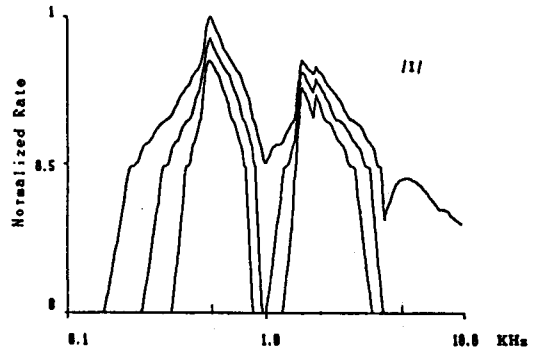
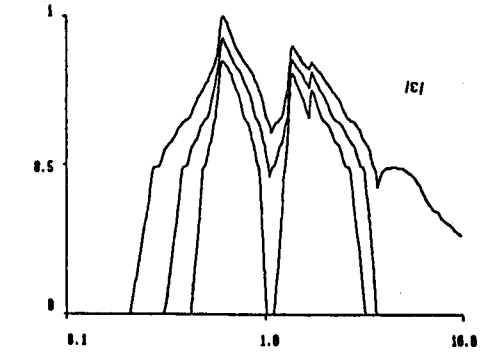
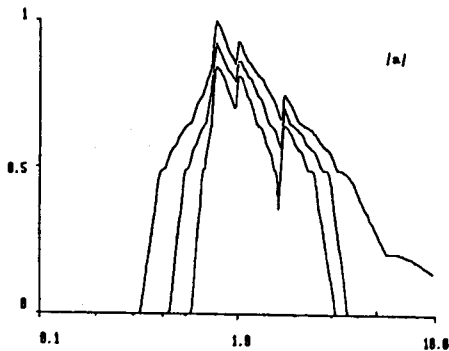


그림 4-3 40,50,60dB에 대한 청각신경의 발화패턴

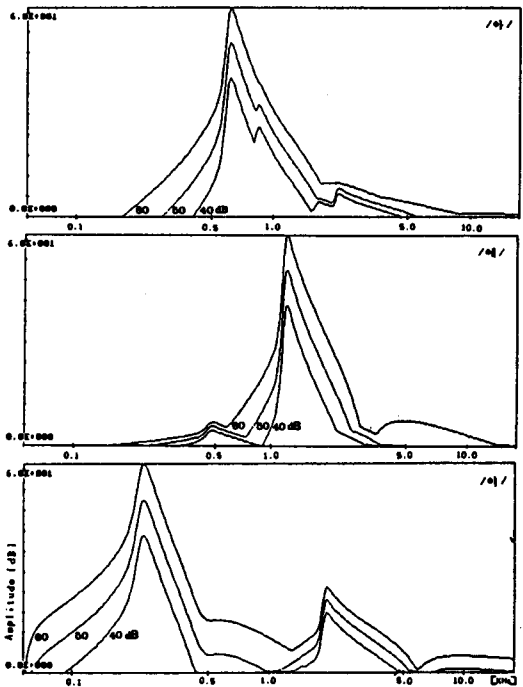


그림 4-4 모음에 대한 자극패턴

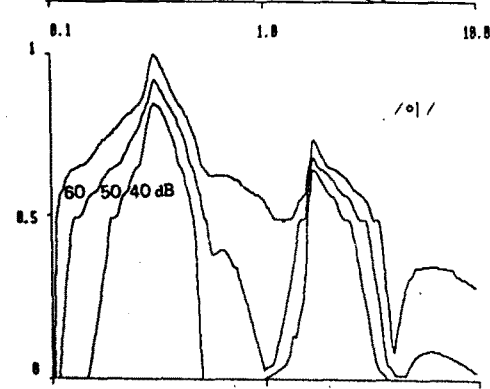
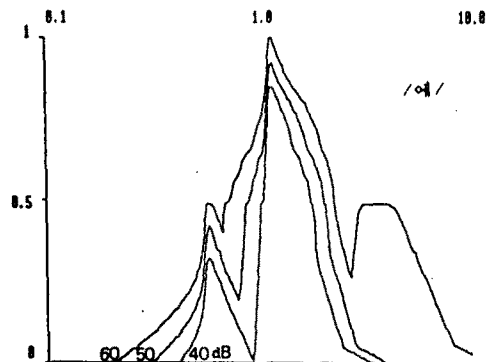
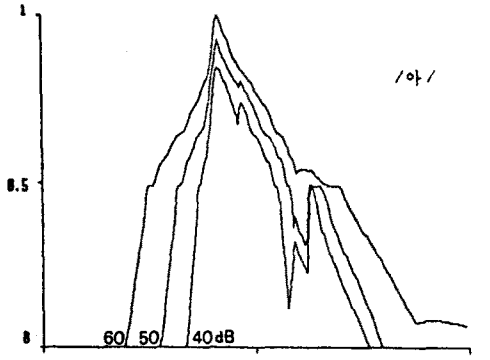


그림 4-5 40, 50, 60dB 자극에 대한 발화 패턴

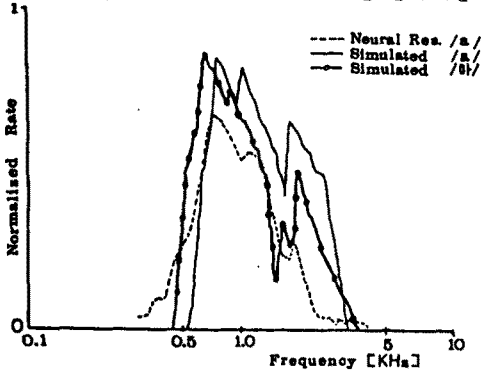


그림 4-6 /a/와 /a/에 대한 응답 특성

그림 4-6에서 점선은 모음 /a/에 대한 실험 곡선으로서, 입력자극의 크기는 37dB이며 실선은 모음 /a/에 대한 스펙트럼과 신경동조 모델로부터 얻어진 40dB 자극에 의한 발화 패턴이다. 모음 /a/에 대한 곡선은 모음 /a/의 스펙트럼과 신경동조 모델에서 얻어진 40

dB 자극에 대한 발화 패턴이다. 여기서, /a/에 대한 발화 패턴은 음성신호에서의 포먼트 주파수 부근에서 최대 발화율을 나타내며, /a/에 대한 발화 패턴도 포먼트 주파수의 크기에 따른 최대 발화율을 나타낸다. 따라서, /a/와 /a/는 각각 포먼트 주파수 근처에서 최대 발화율을 나타내며, 그의 발화 패턴은 실측치와 유사하다. 또한, /a/와 /a/가 최대 발화율을 나타내는 주파수가 약간 다른 것은 조음(articulation)상의 차이에 기인한다.

5. 결론

본 연구에서는 내이의 손상에 의하여 청각 장애가 있는 사람에게 음성에 대한 지각을 일으킬 수 있는 적당한 자극 패턴을 추출하며, 내이에서의 음성 정보 처리 과정을 시뮬레이션 하였다.

따라서, 본 연구에서는 내이의 생리학 특성을 만족 하는 신경동조 모델과 음성에서 추출된 정보에 의하여 청각 신경에 입력되는 자극 패턴을 추출하였으며, 이 자극 패턴을 청각 신경 모델에 입력하여 발화 패턴을 추정함으로써 음성에 대한 지각을 일으킬 수 있는 가능성을 시뮬레이션 하였다.

1. 음성의 모음에서 추출된 포먼트 주파수와 크기는 음성 인식의 중요한 정보를 갖는다. 즉, 포먼트 주파수는 자극되는 청각신경의 위치를, 그리고 그 크기는 청각신경의 발화율을 나타낸다.
2. 신경동조 모델은 기저막 상의 각 위치에서 특성 주파수에 잘 반응하는 생리학적 특성을 만족한다. 즉, 저주파수는 헬리코트레아에서, 그리고 고주파수는 스테이프스 근처에서 동조가 일어난다.
3. 제1, 제2, 제3 포먼트에서 최고 발화율을 나타내는 청각신경 모델에서의 응답 패턴은 Young과 Sachs 가 실험한 신경 응답 패턴과 유사함을 보였다.
4. 신경동조 모델에서 얻어진 자극 패턴은 음성신호에 대한 지각을 일으킬 수 있으며, 문헌에 의한 음성 인식의 가능성을 보였다.

이상과 같은 결과로부터 기저막 상의 청각신경과 전극의 위치에 대한 상세한 Interface를 고찰함으로써 뇌에서 지각을 일으키는 정보를 청각신경에 주입할 수 있을 것으로 기대된다.

6. 참고 문헌

- 1) G. von Békésy, Experiments in Hearing, Robert E. Krieger Publishing Company, Huntington, New York, 1980.
- 2) N.Y.S. Kiang and E.C. Moxon, "Tails of tuning curves of auditory nerve fibers," J. Acoust. Soc. Am., 55, pp. 626-638, 1974.
- 3) S. Seneff, Pitch and Spectral Analysis of Speech Based on an Auditory Synchrony Model, Ph.D. Dissertation, MIT., 1985.
- 4) S.T. Neely and D.O. Kim, "A model for active elements in cochlea biomechanics," JASA, 79(5), PP. 1472-1488, 1986.
- 5) J.L. Flanagan, Speech Analysis/Synthesis and Perception, Springer-Verlag, 1972.
- 6) E.D. Young and M.B. Sachs, "Representation of steady-state vowels in the temporal aspects of the discharge patterns of population of auditory nerve fibers," JASA, 66(5), PP. 1381-1403, 1979.
- 7) M.M. Sondhi, "Method to computing motion in a two-dimensional cochlear model," J. Acoust. Soc. Am., 63, pp. 1468-1477, 1977.
- 8) S.H. Park, S.H. Beack, and S.K. Yoo, "Neural tuning model analysis of auditory function," 12th International Congress on Acoustics, Proceeding, Vol./Band 1 B3-3, Toronto, Canada, 1986.
- 9) 박상희, 백승화, 유선국, "코클리어 기저막 운동의 2차원 모델 해석," 대한의공학회지, 5권, 제2호, 1984.