COCHLEAR PROSTHESES 405
Reprinted from
ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

COMPUTERIZED TESTING OF SIGNAL-ENCODING STRATEGIES WITH ROUND-WINDOW IMPLANTS [a]

N. Dillier, T. Spillmann, and J. Güntensperger

ENT Department
University Hospital
CH-8091 Zurich, Switzerland

INTRODUCTION

Our team has started work on cochlear implants in 1974. Our first implant operation was performed in 1977 in a 50-year-old patient deaf from meningitis since childhood. Two bipolar platinum electrodes, placed in the modiolus, could be directly accessed through a biocarbon percutaneous plug. Psychophysical testing during a 6-month period revealed frequency difference limens comparable to those of normal-hearing persons below 300 Hz.[3]

Because our experiences with the hard-wired percutaneous connectors had not been very convincing (the plug was inadvertently removed after 6 months), we decided to use only transcutaneous inductive coupling systems in the future. Results with temporarily placed removable electrodes at the round window in five patients in whom reproducible electroneural hearing sensations could be evoked as well as reports from other groups using round-window stimulation [1,5] encouraged us, however, to try chronically implanted round-window electrodes. The results will be compared later with our modiolar electrode data and with other single-channel intracochlear stimulation data.

MATERIALS AND METHODS

We consider these electrode implantations nevertheless to be experimental surgical procedures and therefore require informed consent and total bilateral peripheral hearing loss that cannot be treated by conventional surgical or prosthetic means. Our patients should also be willing to cooperate in our clinical evaluation experiments after implantation.

Since the time of the aforementioned single percutaneous two-channel modiolar implant experiment, another four totally deaf patients have been implanted. Pretests by ear-canal electrodes (reported elsewhere [3]) have shown positive results. TABLE 1 lists the clinical information. The patients had been informed about the experimental nature of this procedure several months before the operation. The first patient, R.G. (born in 1956), was congenitally deaf in both ears. The second patient, E.P. (born in 1950), had bilateral loss of cochleovestibular function after meningitis at the age of 9 months. The third patient, U.T. (born in 1941), experienced sudden hearing loss 2 years before implantation. The fourth patient, C.A. (born in 1961), also had bilateral loss of cochleovestibular function after meningitis at the age of 14. In all

[a] This work was supported by Grant 3.848.0.79 from the Swiss National Research Foundation.

0077-8923/83/0405-0360 $1.75/0 © 1983, NYAS

patients, no useful acoustic information was obtained by a conventional acoustic hearing aid. Data on only the first three patients are presented here since too little time has elapsed since the fourth was operated on.

The technical specifications of our implant have already been described elsewhere.[10] The active ball electrode (approximately 1 mm in diameter) and the reference electrode are made of platinum/iridium, and the receiver coil and demodulator circuit are encapsulated in medical-grade epoxy. The external part of our prosthesis consists of the microphone, preamplifier, and signal-processing circuits (all contained in a box measuring $10 \times 6 \times 3$ cm) and the transmitting coil, which is placed behind the auricle, attached to an ear mold.

The surgical intervention consists of use of a typical access to the middle ear to perform mastoidectomy and electrode placement via facial recess and fixation with fibrin adhesive.

Testing is performed by means of a computerized system. Hearing and discomfort thresholds are determined in every test session. The responses to nonverbal frequency and amplitude differences, and pitch and intensity scaling, are systematically investigated and can be used to compare different electrode systems available in the future.

Speech material, such as numbers, monosyllables, spondees, and sentences, can be stored and analyzed with the same system. Analysis parameters (pitch, gain, formants, vocal-tract filter coefficients, and zero-crossing intervals) are used to generate on-line stimulation signals according to several processing algorithms. Stimulus parameters are fully programmable. Complex stimuli such as speech can also be processed and delivered by the computer.[4]

Results

The first measurements after the implantation were repeated determinations of the threshold of hearing and discomfort at several frequencies. Thresholds remained stable over the whole experimental period of 14 (R.G.), 10 (E.P.), or 3 (U.T.) months (Fig. 1). The threshold of discomfort was between 8 to 12 dB above the hearing threshold. Discomfort thresholds were measured as quickly as possible, and were therefore not as accurate as hearing threshold measurements.

Table 1

Summary of Information on Patients with Round-Window Implants

Patient	Age (yr) & Sex	Age at Onset of Deafness (yr)	Years Deaf:	Cause of Hearing Loss
R.G.	25, F	0	25	Congenital
E.P.	31, M	0	31	Meningitis
U.T.	41, F	39	2	Sudden deafness (left); congenital (right)
C.A.	20, F	14	6	Meningitis

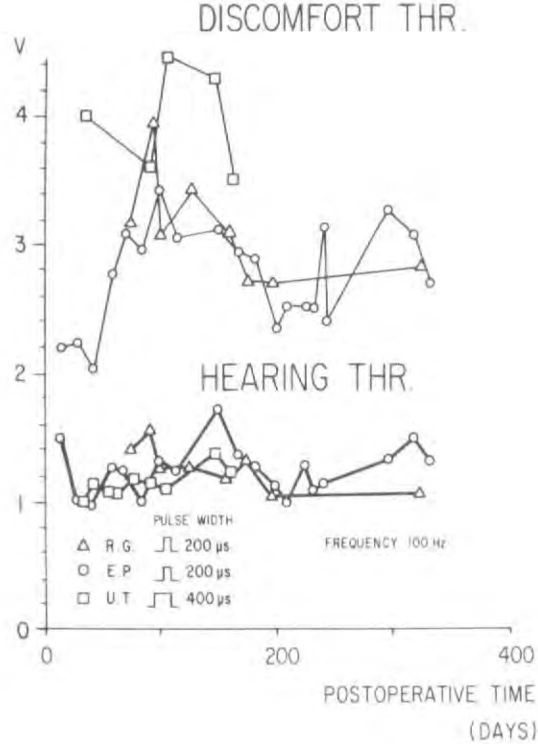

FIGURE 1. Thresholds of hearing and discomfort over time. THR, threshold; V, pulse amplitude (in volts) at transmitter-stage input of portable sound-processor.

Frequency difference limens could be determined up to 500 Hz. They were in the same range as the values of our first patient with modiolar electrodes.[4] FIGURE 2 shows two examples of pitch scaling. In the first experiment the patient was asked to scale frequencies between 20 and 100 Hz, in the second between 100 and 1000 Hz. Every experiment consisted of 100 trials with ten linearly spaced frequencies. Every frequency was repeated ten times in random order. The mean of these ten scalings is displayed in FIGURE 2 in logarithmic frequency scale. It can be seen that scaling is possible up to approximately 500 Hz. Frequencies below 100 Hz seem to be scaled slightly differently from frequencies above 100 Hz.

In a series of scaling experiments, the effect of training and feedback was investigated. FIGURE 3 shows normalized standard deviations plotted against the session numbers. Triangles denote experiments without feedback. Circles denote experiments with feedback, that is, the patient was shown the stimulated frequency value (in percent of chosen interval) after every trial. Training considerably improves the reproducibility in both conditions, and feedback will lead to even more consistent responses. All these experiments were carried out on one morning, with only short interruptions between each session. Considering the size of the scatter, it cannot be excluded that different feedback conditions

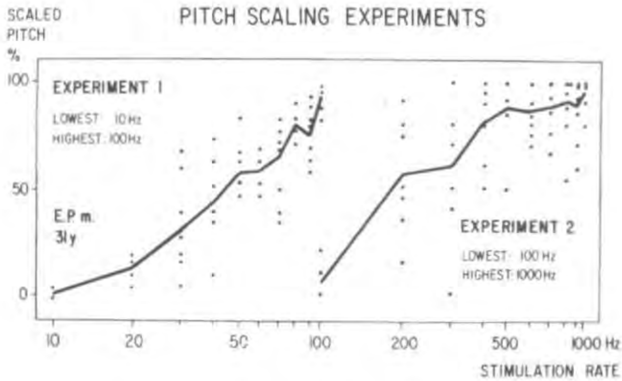

FIGURE 2. Pitch scaling examples in (1) low-frequency range and (2) high-frequency range.

(such as logarithmic instead of linear frequency scales) might have led to different shapes of the scaling function.

Amplitude difference limen results are shown in FIGURE 4. The values of patient E.P. were much lower and practically independent of stimulus level in contrast to the values of R.G. The dynamic range of E.P. was lower (only 8 dB or less) than that of R.G. (10 to 12 dB), which might explain this difference.

Numbers between 13 and 99 were used in identification tests. Every test comprised 40 randomly choosen numbers. FIGURE 5 shows the discrimination results over time for R.G. with the linear regression line (the correlation coefficient is statistically significant at $p < 0.001$). The tests were carried out with the portable speech-processor and the same speaker in all experiments. FIGURE 6 shows comparisons of speech test results in three conditions (S: stimulation alone, L: lipreading alone, SL: stimulation with lipreading). Again the

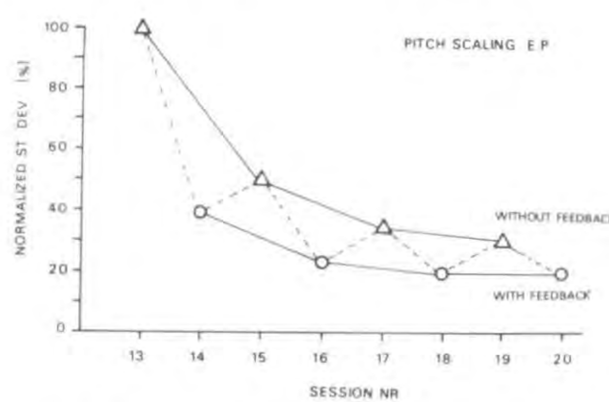

FIGURE 3. Effect of training and feedback in pitch scaling experiments in subject E.P. Every point represents 100 trials in the frequency range of 20 to 100 Hz.

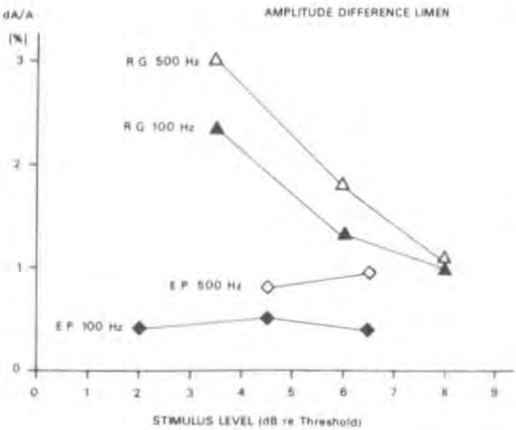

FIGURE 4. Amplitude difference limens. Intensities are referred to individual subjective thresholds.

stimulation tests were done with the portable stimulator and with the same speaker as in the lipreading or in the combined tests. The data were averaged over several sessions.

TABLE 2 lists the number of trials and the relative improvements in discrimination with lipreading compared to stimulation alone. The best effect of electrical stimulation is revealed in the speech-tracking test, where relative improvements of combined acoustic and visual input over visual input alone of 9–125% are obtained. Patient 3, with late-onset hearing loss, who is an excellent lipreader (because of pre-existing moderate to severe hearing loss and special-school education) scored best both with and without a prosthesis. She uses her prosthesis mostly in difficult communication situations where there are many competing speakers. The other verbal tests, voice (speaker sex) identification and the number recognition test, are less helpful in ascertaining specific

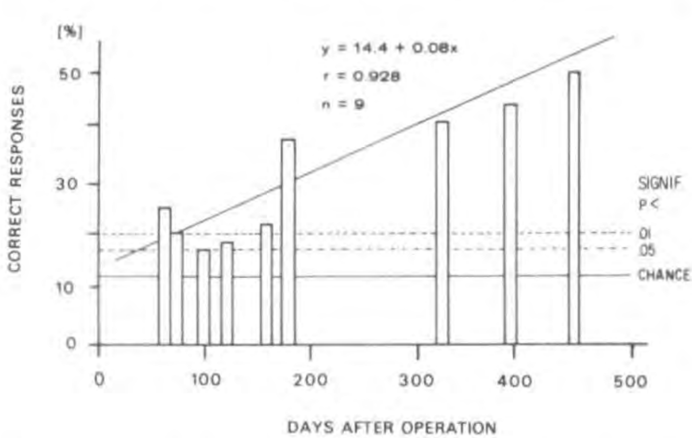

FIGURE 5. Improvement of number identification results with time in subject R.G. (closed test). Every column represents 40 responses to two-digit numbers (13–99) spoken by a male speaker via the portable sound-processor.

responses to the prosthesis either because these tests give extremely high values with electrical stimulation alone (in the voice identification test) or because lipreading alone is nearly perfect (in the numbers test). In the latter case, the addition of the acoustic input to the visual input alone does not provide a substantial gain. The failure of some tests to show a prosthetic gain does not agree with the real-life benefit and leads us to conclude that the choice of an appropriate test paradigm is crucial in electrode evaluation.

For vowel discrimination experiments, a closed set of digitally stored speech sounds has been used. Two amplitude coding options (lin, log) and three pulse-rate coding options (F0 = fundamental, F2 = second formant frequency divided by 18, ZC = zerocrossing intervals) are available with the laboratory speech-processor. The original fundamental and formant parameters of the half-second samples are listed in TABLE 3. Fundamental frequencies are similar for all five vowels, whereas F2 is low for ⟨A⟩, ⟨O⟩, ⟨U⟩ and high for ⟨E⟩ and ⟨I⟩.

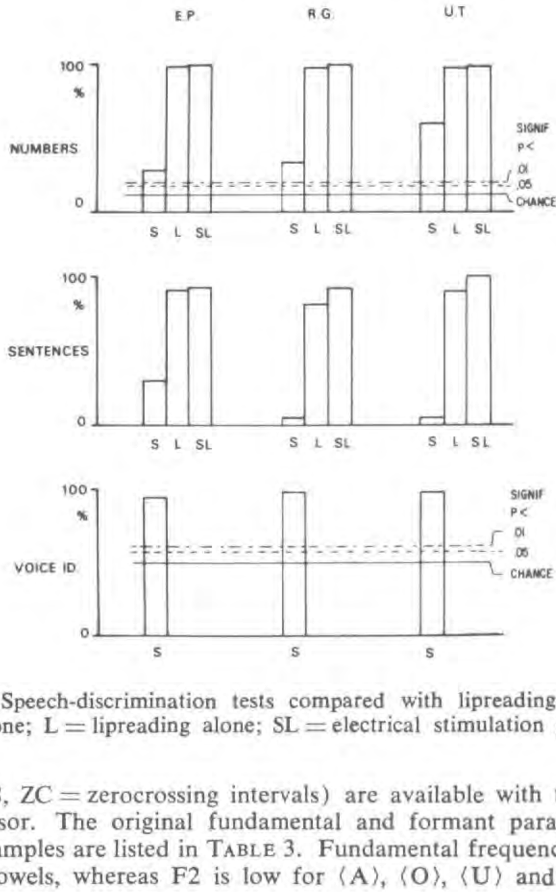

FIGURE 6. Speech-discrimination tests compared with lipreading. S = electrical stimulation alone; L = lipreading alone; SL = electrical stimulation plus lipreading.

Therefore, it can be expected that an F0 algorithm will not work well in separating these sounds, whereas an F2 algorithm should discriminate between the two groups of low (A, O, U) and high (E, I) second formant frequency. This is documented in TABLE 4, which shows confusion matrices for both stimulation algorithms. The contingency coefficient (which would be 1 if no confusions had occurred) is not significant at the 5% level for the F0 algorithm.

TABLE 2

OVERVIEW OF SPEECH-DISCRIMINATION RESULTS WITH ROUND-WINDOW ELECTRICAL STIMULATION

	Voice Identification	Number-Recognition Test			Speech Tracking		
Patient	SA (%)	SA (%)	LRA (%)	S + LR (%)	LRA (wpm)	S + LR (wpm)	Relative Improvement (%)
R.G.	98	33	97	99	12.1	13.2	9
E.P.	95	28	98	99	17.8	23.3	31
U.T.	98	60	97	98	19.2	43.2	125

ABBREVIATIONS: SA = discrimination by stimulation alone; LRA = discrimination by lipreading alone; S + LR = discrimination by combined stimulation plus lipreading; wpm = words per minute.

NOTE: Relative improvement $\frac{S + LR - LRA}{LRA} 100\%$.

The matrix for the F2 algorithm shows a contingency coefficient of 0.88 ($p < 0.001$). If only the two subgroups (A, O, U and E, I) are considered, this coefficient rises to 0.95.

However, vowel identification is probably not the main objective of a cochlear prosthesis. Acceptability and objective advantage may well depend on other factors. In a series of experiments, some speech-processing strategies were tried. FIGURE 7 shows a comparison of subjective scores and discrimination results in numbers tests. Only the four or five algorithms that scored best were tested systematically. Again substantial patient differences are apparent. Discrimination results and subjective scores correlate reasonably well for patients U.T. and R.G. It is interesting to note that in all patients results with the tests with the portable sound-processor (live speech) showed the best range, whereas the zerocrossing algorithm of the laboratory stimulator attained the same result as that of the portable stimulator only in patient E.P. The most probable explanation for this outcome seems to be the test situation. Another reason could be that the amplitude coding of the portable stimulator is neither linear nor fully logarithmic and is not exactly reproduced by the computer

TABLE 3

VOWEL IDENTIFICATION TESTS: STIMULATION PARAMETERS

Vowel	F0 (Hz)	F1 (Hz)	F2 (Hz)
A	153	724	1044
O	165	349	621
U	177	270	681
E	160	330	2467
I	183	198	2586

NOTE: Pulse-rate coding: F0, F2/18; amplitude coding: linear, logarithmic.

algorithms. From FIGURE 7 it also seems that pulse-rate coding by the fundamental frequency is superior to F2 coding. This might be even more true for longer speech samples, such as sentences. There is no conclusive evidence yet to allow us to decide whether coding of unvoiced (consonantal) speech sounds by high (over 200 Hz) or low (below 200 Hz) random pulse sequences is to be preferred.

TABLE 4

VOWEL CONFUSIONS IN PATIENT R.G. WITH TWO DIFFERENT ALGORITHMS [a]

	Response					
Stimulation	A	O	U	E	I	n
A	8	3	4	2	3	20
O	1	1	6	6	6	20
U	0	4	4	3	8	19
E	13	5	1	1	0	20
I	3	4	3	3	6	19
n	25	17	18	15	23	98

Rate: F0
Amplitude: linear
20.4% correct
Contingency coefficient = 0.44
NS

	Response					
Stimulation	A	O	U	E	I	n
A	16	1	1	0	0	18
O	4	11	4	0	1	20
U	6	3	10	0	1	20
E	0	0	0	18	2	20
I	2	0	0	6	12	20
n	28	15	15	24	16	98

Rate: F2
Amplitude: linear
68.4% correct
Contingency coefficient = 0.88
$p < 0.001$

[a] Pulse rate determined by fundamental frequency and pulse rate determined by second-formant frequency (divided by 18).

DISCUSSION

The results obtained with implanted electrodes at the round window are comparable to results with electrodes in the scala tympani or in the hearing nerve with respect to frequency resolution. Although the subjective quality of the perceived sounds is expected to vary depending on the stimulation site,[1, 2, 5, 9] it seems that timing effects of electrical stimulation are the principal factors of differential pitch perceptions.

The dynamic range with round-window stimulation seems to be somewhat lower than that with intracochlear stimulation, although the highest range (12 dB) is within the range obtained with scalar electrodes. Sinusoidal stimulation may give a slightly higher dynamic range, especially at lower frequencies, but we prefer to use pulsatile stimuli for safety as well as technical reasons.

Whether the differences in speech discrimination between our results and those of other groups[7,8] using single-channel access can be explained by the place of stimulation remains to be shown in further studies. In contrast to the assertions of some authors using intracochlear single-channel electrodes, we could not achieve true open-word discrimination in any of our patients with any of the tested algorithms. Differences of testing methods and of the amount of training of implanted persons might play a substantial role in the variability

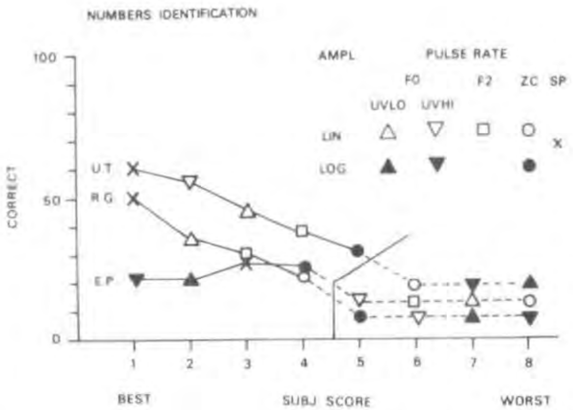

FIGURE 7. Speech-encoding algorithms: correlations between subjective scoring and discrimination results: two amplitude coding options (linear, logarithmic), pulse-rate coding by fundamental frequency (F0), second formant frequency (F2) divided by 18, and zerocrossing intervals (ZC). During unvoiced periods either a low (UVLO) or a high (UVHI) random pulse is generated. The portable sound-processor (SP) uses zerocrossings and automatic gain control as well as logarithmic compression, which is adjusted individually for every patient.

between implant centers. The small group of patients and too many clinical variables (cause, age at onset, and duration of deafness) prevents further speculation on preferred coding formulas at present. The patient with the best test results (U.T.) is postlingually deafened; the other two are prelingually deaf. Further improvement over time is still possible, as shown for results in number recognition for R.G. over 15 months (FIG. 5). The same might be the case for speech tracking, which in our experience is the best single test to obtain real-life benefit scores of electrical stimulation. Still, there remain other concepts of speech encoding not tested yet which may improve discrimination.

Our technique of electrical stimulation of hearing is a safe procedure involving only routine middle-ear surgery to place the induction coil within the mastoid bone and to put an electrode at the round-window margin. By an external stimulator, sounds are transformed to electrical stimuli suited to excite surviving fibers of the cochlear nerve. This helps the patient to communicate and provides some basic cues for speech discrimination. The ability to profit from these percepts can be considerably improved by accustomation and training, which are still in progress in all our patients.

SUMMARY

After extensive testing of a patient with two bipolar modiolar electrodes connected to a percutaneous plug in 1977, we provided four patients with a single-channel monopolar round-window electrode connected to a tuned radio-frequency receiver coil. Loudness and pitch discrimination and results of psychophysical scaling experiments of extracochlear electrodes are comparable to those with intracochlear stimulation. Extensive testing with a computerized test system and with tape-recorded and live speech material showed that accurate vowel and speaker identification by stimulation alone is possible and that discrimination by lipreading is considerably improved. Interactive training sessions further improve discrimination results. Different signal-encoding algorithms can be used to generate in real time stimulation signals from prestored speech parameters (such as pitch, gain, formants, and zerocrossing intervals).

REFERENCES

1. CHOUARD, C. H. & P. M. MACLEOD. 1976. Implantation of multiple intracochlear electrodes for rehabilitation of total deafness: Preliminary report. Laryngoscope **86:** 1743–1751.
2. CLARK, G. M., Y. C. TONG, L. F. A. MARTIN & P. A. BUSBY. 1981. A multiple-channel cochlear implant. Acta Otolaryngol. **91:** 173–175.
3. DILLIER, N., T. SPILLMANN, U. P. FISCH & L. J. LEIFER. 1979. Encoding and decoding of auditory signals in relation to human speech and its application to human cochlear implants. Audiology **19:** 146–163.
4. DILLIER, N., J. GÜNTENSPERGER & T. SPILLMANN. 1980. A computer-controlled test system for electrical stimulation of the auditory nerve of deaf patients with implanted multielectrodes. Scand. Audiol. Suppl. **11:** 163–170.
5. EDDINGTON, D. K. 1980. Speech discrimination in deaf subjects with cochlear implants. J. Acoust. Soc. Am. **68:** 885–891.
6. FOURCIN, A. J., S. M. ROSEN, B. C. J. MOORE, E. E. DOUEK, G. P. CLARK, H. DODSON & L. H. BANNISTER. 1979. External electrical stimulation of the cochlea: Clinical, psychological, speech-perceptual and histological findings. Br. J. Audiol. **13:** 85–107.
7. HOCHMAIR, I. J., E. S. HOCHMAIR, R. E. FISCHER & K. BURIAN. 1980. Cochlear prostheses in use: Recent speech comprehension results. Arch. Oto-Rhino-Laryngol. **229**(2): 81–98.
8. HOUSE, W. F. 1976. Cochlear implants. Ann. Otol. 85 (Suppl. 27): 1–92.
9. SIMMONS, F. B. 1966. Electrical stimulation of the auditory nerve in man. Arch. Otolaryngol. **84:** 22–76.
10. SPILLMANN, T., N. DILLIER & J. GÜNTENSPERGER. 1982. Electrical stimulation of hearing by implanted cochlear electrodes in humans. Appl. Neurophysiol. **45:** 32–37.