From the linguistic point of view the Modern Northern English dialects include a complex range of vowels and diphthongs, which are not generally used in Standard English. The range of diphthongs is very wide. Quite considerable differences in dialect will be found between the speech of inhabitants of villages separated by only a few miles, and the field of investigation is very extensive. It gives me great pleasure to state that we have always found our speakers particularly anxious to help and I should like to thank them for their courtesy.

45. Prof. G. OSCAR RUSSELL (Ohio): *Synchronized X-ray, oscillograph, sound and movie experiments, showing the fallacy of vowel triangle and open-closed theories.*

It is to be hoped that all of you were fortunate enough to see the very fine talking X-ray movie of Prof. MENZERATH. You noticed how movements of the velum, larynx, and tongue could be readily followed. And that the tongue was not the only organ occupied in creating vowel quality differences. Furthermore, that the back throat was obviously fully as much, if not more, involved than the front mouth.

So again we have ample proof that the unfortunate physiological front mouth vowel triangle does not represent the facts. This opportunity should not be permitted to pass, therefore, without calling attention to the necessity of the linguistic scholar and teacher adopting a more reliable classification scheme.

Since it must still be said we know practically nothing as to the physiological cause of fine vowel distinctions, and certainly that the physiological act is far more complex than the mere front tongue arching represented by that vowel triangle, the folly of holding to, and fallacy of the latter, must be obvious. But what can be substituted? That is the inevitable question. And a proper one. Let us consider it a moment.

Sound change is dependent on what we hear. Not on what our tongues feel. And the *normal learning process is also guided by the auditory* sense rather than by what the tongue feels. Where we have to rely on the latter even when supplemented by the visual, as in teaching the deaf, the process is so slow, and the results so inaccurate, as to make that very apparent. Since what we hear is obviously primary, our classification schemes, transcription characters, and terminology, should also be acoustic. For it is just a wildly unscientific absurdity to listen to a speaker's strange sound, and then proceed to record it on paper in terms of a physiological character and scheme. Especially where the latter are fantastic, and unsupported by every X-ray and other scientific experiment we carry out; being based, as they are, solely on a physiology which was originally, and is still, purely imaginary.

It should of course be recognized by all that the ear will not hear many fine distinctions in sound, which a high speed scientific experiment would record. The same thing holds true of the eye. It cannot see the flight of a bullet; hence if you want to know just what happens, a high-speed motion picture of from 3000 to 90,000 exposures per second is more reliable. And it will tell more than the most voluminous argument about, and description of, what one, or a group, imagines the eye can, or should, see. But in that realm we understand the manifestations shown on the experiment. Whereas when we look at the experimental analyses of the vowel we, as yet, understand but little of what we see. The phenomena are so complex, and are dependent on so many misunderstood processes, that we stand baffled before our experiment once we have made it.

Where this is not true there are of course many details we can establish with far more reliability than any number of listening recorders could do depending solely on their ears. Is the vowel nasal, for example? Or is it partially unvoiced? Does it start before the consonant ends? And so on *ad infinitum.* It is sheer folly, in such cases, to reject the aid careful scientific experiment could lend, and to depend solely on a series of letters to record what the ear hears. That would be like the chemist who in this day and age attempted to write a treatise on acids and their physiological reactions, by recording in chemical and other learned symbols just what his tongue tasted and other senses perceived. That day is long since gone. And real scholars and scientists would in this day and age laugh such a procedure to scorn. For the available technique makes far greater accuracy possible.

Generally speaking we can say that a careful scientific analysis of all consonant manifestations is as much called for as in the case of the chemical problem above. For we do know what we see on those experiments. Is it a stop, fricative, velar, partially voiced or nasal? etc. He who relies on his ear when he can so much more accurately establish the facts by simple scientific experiments is to-day as far behind the times as the above-mentioned antiquated physiological chemist.

We also regularly detect that a given e is more "closed" than some ɪ. Unfortunately, though, that is exactly what the linguistic scholar does not want us to find. For he is thinking in terms of what he hears and describing in terms of a physiology he imagines, which in actuality is wide of the mark. Obviously, then, the wise thing for him, and you, and me, is to cease that scientifically absurd process and describe, as well as think, in terms of what we all three hear. Then when we classify, use likewise a scheme which is based on the acoustic, rather than the antiquated imaginary physiological manifestations which are actually non-existent.

The cardinal vowel device of Prof. DANIEL JONES is to that end recommendable. The phonograph record makes it possible for anybody to compare the vowel being considered with the reference norms given thereon. We have made such phenomenal progress in phonograph recording during the last decade that such an acoustic "yard-stick" is now quite reliable. Of course it should be re-recorded at intervals so as to keep it up to date as recording techniques are improved (assuming of course that the old are always kept available for verification of pronunciation uniformity). And it is my personal conviction that the original matrix of the stampers should be authorized by such a congress as this and officially left in central authoritative places, as is the metre in Paris, or yard-stick norm at the Bureau of Standards in Washington.

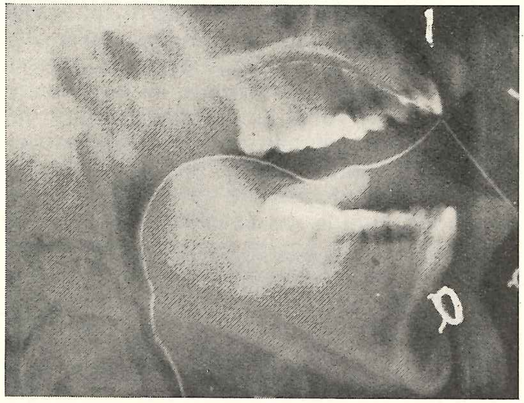

Fig. 1. Prolonged vowel ɑ where the subject is told to "open your mouth and prolong the ɑ in *cock* until I tell you to stop".

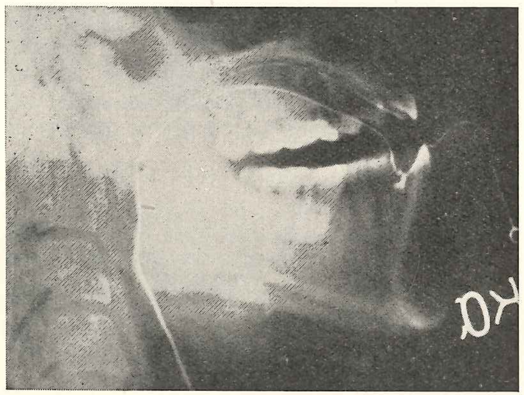

Fig. 2. Vowel ɑ in *cock* of the same Mid-West American subject as Fig. 1, subject allowed to speak normally. Caught in the sentence, *don't* **cock** *that gun it might go off*. X-ray exposure of 1/120 of a second.

We might have some hope, then, that careful scholars, at least, would cease to use terms, schemes and foolish characters or diacritical marks that are based upon imaginary physiological causes (such as *open* and *closed*) which the X-ray has proved to be not true and contrary to fact.

For Roentgen studies under widely varying techniques, and in

many different countries and by numerous different investigators, are so definitely confirmatory as to leave no room for doubt as to many of those physiological facts such as the first above mentioned.

They show conclusively that, for example, the cavity of the mouth may be fully as closed for ɛ as for ɑ. Cf. Figs. 1, 2, 3 and 4, and

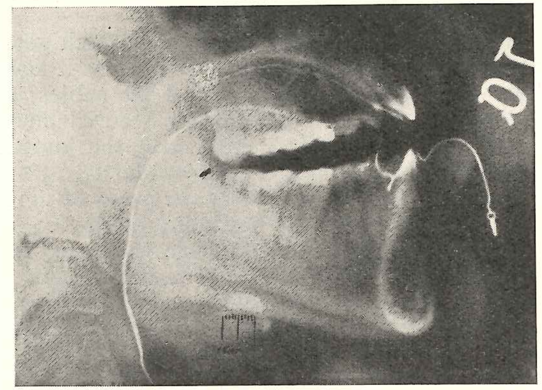

Fig. 3. Vowel ɑ in *tot*. Same subject. Ordinary discourse. The part herewith is taken from the sentence, *poor little* **tot** *is almost asleep right now*.

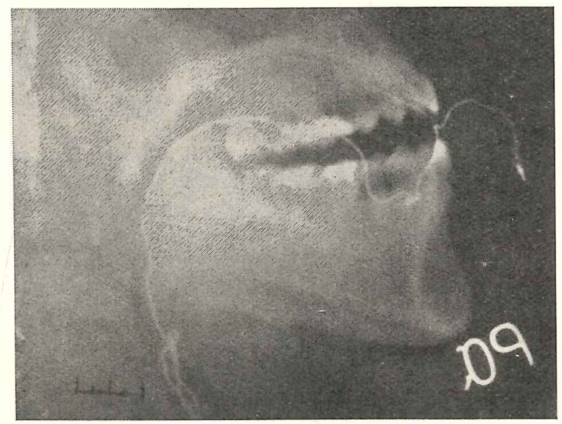

Fig. 4. Vowel ɑ in *pop*. Same subject. Same normal discourse and exposure speed. From part where sentence occurs, *and I'll* **pop** *him right on the head*.

you will see they also show that many different front tongue positions may be taken even by the same subject to produce the same vowel. The technique by which the X-ray photos were made which are reproduced in these figures presents us with the exact facts of normal speech. So there can no longer be any question as to the validity of this conclusion.

They are made of the normal speech of a subject talking just as I am talking now. And the exposure makes itself. In other words it takes place automatically on the vowel predetermined by the investigator, and at the precise point in the course thereof previously fixed as desirable. Fig. 5, which is part of the synchronized whole, tells you exactly where that was and may help clarify for the reader what is meant.

The top line is made by a recording needle which goes down the instant the vowel starts and stays down as long as it lasts.

The bottom line is made by the voice of the subject actuating a timer which makes a peak every 1/120 of a second. They begin the instant the vowel starts and last only until the instant it stops. In this case the vowel was the ʌ of *pup* and lasted 13/120 (or 1/10) of a second, as can be seen if one counts the peaks of the lower line.

Drop a perpendicular from the beginning of the exposure dip in the upper line to the corresponding peak in the lower line, and you see the X-ray started 4/120 second after the vowel got under way. Another will show you it lasted 1/120 second. In other words you know now just exactly how long the exposure lasted, and precisely at what point in the course of the vowel it occurred.

The apparatus and technique make it possible to vary that point at will. If you want the exposure to take place 7/120 instead of 4/120 of a second after the vowel starts you can so predetermine. Or if one wants it at the end or the beginning of the vowel, he may so fix it.

An examination of the lower line shows, further, that the subject continued talking in his normal, undisturbed speech, and no other exposures took place on any other vowel. You obtain your X-ray picture of the vowel you want at precisely the point in its articulation where you want it.

This is the reason we discarded a motion picture X-ray technique, in which field we were pioneers[1] now well on to two decades ago, and we chose to develop this one in order to accomplish some things the former would not do. In the first place the motion picture exposure is too slow, since it is as yet impossible to move at the rate of 120 exposures a second as is here shown would be necessary. One is limited to from 10 to 16. As you will see from Fig. 6 (where each thin perpendicular line represents 1/1000 second and each broad one equals 1/100 of a second), many sounds such as this vowel ə last only 3/100 second.

Hence all the motion picture exposures at the rate of 16 per second would skip right over this vowel. 100 per second as in our technique would give you but 2 exposures and possibly only 1.

The most serious objection to the motion picture far too slow exposure technique lay, however, in the fact that one thereby takes a sampling of views which fall at uncontrollable indiscriminate places in the speech record. The first frame may occur right as the tongue is starting to move into position for a sound. The next long after that sound is finished, and in the middle of the following one.

[1] See *The Vowel.*

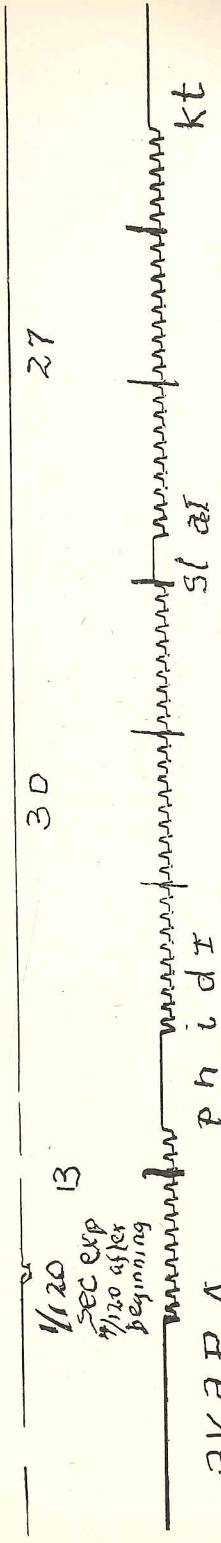

Fig. 5. Extract from part of the synchronized record. The upper line records the X-ray exposure: instant it started, and duration. Bottom line: time, each peak being 1/120 of a second; instant the vowel started, being first peak, and ended as shown by last peak; hence also its duration. The straight lines occur on consonants and show their duration.

As the subject was talking along, he came to the sentence here extracted, saying ...*of a pup he disliked.* The exposure took place on the ʌ of *pup* 4/120 of a second after the vowel started, lasted 1/120 of a second, and the vowel continued 8/120 of a second longer. All the vowels hereon were excessively long since he was stressing each of the last three words as the meaning demanded.

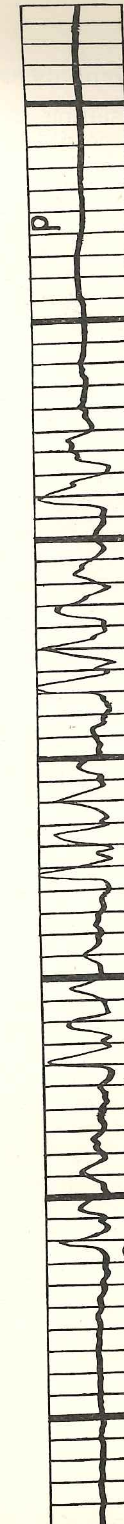

Fig. 6. Extract from the oscillographic record. The ə written *a* in the *of a pup he disliked.* The thin lines occur every 1/1000 of a second. The thicker lines every 1/100 second. As will be seen this vowel was only slightly more than 3/100 of a second in duration before the p began.

So while the resulting record may serve to make a fairly reliable study of motion possible, it would be "a delusion and snare" in any attempt to establish the facts as to the cause of vowel quality differences. Hence the excessive additional expense hinders rather than helps.

Fig. 6 is from an oscillographic tracing which we synchronize with the X-ray picture and the record shown in Fig. 5. Consequently if one predetermines the X-ray exposure to fall in the middle segment, that one can thereupon be submitted to a harmonic analysis, inspection, or other means of determining the physical characteristics of this particular vowel at that instant.

So far as the author can determine, this is the first time such a correlation of the physics and physiology, and such a complete analysis, has ever been made possible. The usual process has been

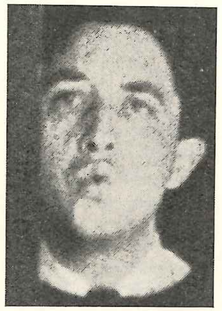

Fig. 7

Fig. 8

Fig. 7. Prolonged vowel **u**. In the same word as Fig. 8. Other conditions also identical.

Extract from the accompanying synchronized motion picture record. It shows the mental attitude of the subject, in this case deadly serious, gestures, facial, and emotional features generally. Also necessary control information such as position and tilt of head, etc.

Fig. 8. Vowel **u** in *pooh poohed.* Normal speech of sentence extracted from connected discourse being . . . *and pooh* **poohed** *those who liked them.*

to make an isolated and independent physical (oscillographic) analysis and thereupon assume the physiological. Or, as the author has done in his preceding studies of the vowel, make the physiological (X-ray), and either assume, or attempt the impossible task of trying to interpret the same in the light of independently and separately made physical analyses. Either process is unscientific, as most of us long ago recognized.

Of course another great failing of all X-rays and physical or oscillographic analyses to date has been that no synchronized phonographic record accompanied the same. Hence when we discovered anything we had not expected, we always felt vaguely uneasy as to whether the vowel sounded exactly as we assumed it did. "Maybe the subject's speech was dialectical. Possibly he slipped. Perhaps his normal speech was disturbed or otherwise distorted."

So under our present technique a synchronized sound record is provided. Likewise a frontal motion picture of the subject's gestures and facial expression. So we can listen back to just how he pronounced at any given point. And we can see any ill at ease, composed, cramped, amused, frightened, hesitant, or other emotion or mannerisms he manifested. An extract from such a section is shown in Fig. 7. So when we get through we have a complete record of the sound.

What are the results? That cannot be completely reported in this short paper, nor in many more to come for some years in the future, as you will all agree. But they do confirm the facts above noted.

They show that if the subject is just told "open your mouth and say *ah*" (ɑ) the tongue position will be entirely unlike that of any vowel of normal speech though heard as being identically the same. That can be seen from Fig. 1 of such a vowel, compared with Fig. 2 of the Mid-West American **ɑ** in *cock*; or the **ɑ** in *tot* of Fig. 3; or the **ɑ** of *pop* in Fig. 4. We might have surmised as much from external motion pictures of the lips. For the same thing holds true of jaw and lip position. That can be seen from Fig. 8, showing the normal ordinary conversational lip position for the Mid-West American **u** of *pooh* in the connected discourse where it occurs in the sentence *and pooh poohed those who liked them.* That is if this Fig. 8 vowel **u** is compared with the Fig. 7 vowel **u**, which latter is of the artificial sound created when you tell your subject: "now just speak a prolonged **u** as in *pooh* for me in your normal natural way." A little comparison of the two shows that what he then produced was anything but normal. For the radical puckering was just not characteristic of his regular speech. As anybody can see, the two lip positions are no more alike than are the tongue positions in the X-rays above.

Is that not ample demonstration of the fact that one must needs permit the subject to speak the vowel in normal contextual sentences of his every-day connected discourse, at his characteristic undistorted naturally rapid rate of speed. And that distortion of the vowel by unnatural prolonging in order that we may take slow exposure X-rays is, to say the least, unreliable and inexcusable.

Summary. Our original observation is now again confirmed, namely, (1) A progressive order of tongue arching as postulated in the vowel triangle is untenable as an explanation of vowel quality differences. (2) The same vowel may be produced by many different tongue positions. (3) The back throat and other parts of the vocal mechanism are fully as much if not more involved than the front mouth. The terms "open" and "closed" or "high" and "low" fail to represent the facts. (4) Phonetic characters for sound transcription, their modifiers, as well as vowel theories, classification schemes, and teaching methods, based on those discredited and misleading assumptions, should be changed to fit the facts.

46. Prof. Th. S. Flatau (Berlin): *Über eine neue Methode der Endostroboskopie des Kehlkopfes.*

Die Endoskopie des Kehlkopfes ist in der von mir begründeten und geleiteten Phonetischen Abteilung der ersten Universitäts-Hals-Ohren- und Nasenklinik unter Passow's Leitung regelmässig geübt und gelehrt worden. Die Vervollkommnung des Instruments durch die geometrische Optik mit Hilfe der Firmen Zeiss und Georg Wolff und unter dem wertvollen Beirat von Professor Ringleb, die ständige Beobachtung und Ausarbeitung der Technik, die etwa mit der Laryngoskopie an Schwierigkeit gleichzustellen ist, bewirkten, dass wir stets eine Anzahl von Patienten für die Vorlesungen zur Verfügung hatten, die ohne örtliche Betäubung endoskopirt werden konnten. War das Instrument eingelegt, so konnten 30–50 Studenten das Bild beobachten. Damals entstand in meiner Abteilung die schöne Arbeit Anthon's über die Beobachtung der Schwingungsfiguren und ihre Beeinflussung durch die elektromechanische Tonbehandlung, wobei die Stimmkranken—meist phonasthenische Sänger und Sprecher jedesmal vor und nach der therapeutischen Sitzung endoskopirt wurden.

Die Endoskopie hat dann einige weitere Verwendungen und Entwicklungen gezeitigt. Ich erwähne die Autoendoskopie, deren Autor Herr Prof. Paconcelli-Calzia sich unter uns befindet; dann die Endostereoskopie, deren erstes hervorragendes Modell von mir 1914 durch die Firma Georg Wolff hergestellt worden ist. Es wurde vor dem Weltkrieg zu einer Ausstellung nach London geschickt und da ist es ein Kriegsopfer geworden—vermisst und nie wieder erschienen. Sodann hatte ich die Freude bei mehreren Gelegenheiten (Prag, Leipzig, Berlin) die ersten endoskopisch gewonnenen Reihenaufnahmen von Tönen und Tonfolgen vorzulegen, die mit verstärkter Beleuchtung und kleinen Schmalfilmkammern von Sängern gewonnen waren.

Wenden wir uns nun dem naheliegenden Problem der Endostroboskopie zu, so möchte ich mit der Vorführung des ersten Apparats aus meiner Sammlung beginnen. Der Sehstrahl hatte den Weg Stimmlippen—Endoskopische Optik—Auge des Untersuchers und wurde am Ocular durch eine rotirende Lochscheibe unterbrochen. Ist die Unterbrechungsfrequenz F_1, die Lochzahl am Scheibenumfang z, und die Umdrehungszahl der Scheibe in der Minute N, so ist $F=N/60.z$. Wenn die Stimmlippen mit der Frequenz F_1 schwingen, so ist bei der Unterbrecherfrequenz F_2 die Schwingungszahl im stroboskopischen Bilde ausgedrückt durch $\Delta F=F_1-F_2$. Bei dem gezeigten Modell wurde mit 3 leicht auswechselbaren Scheiben von 8, 24 und 42 Löchern ein Tonbereich von 50–840 Hertz bestrichen; der Motor war aber durch eine Zentrifugalbremse und ausserdem durch einen Widerstand im Griff des Instruments regelbar, wodurch die Frequenz auf 1200 erhöht werden konnte. Das Gesamtgewicht betrug aber 1,5 kg, was die Handhabung erschwerte und auch die Konstanz der Umlaufzahl war nicht zufriedenstellend. Deshalb wurde in den nächsten Versuchen der Motor wieder aus dem Griff entfernt und die Verbindung durch eine biegsame Welle hergestellt. Dadurch wurde die Drehzahl äusserst konstant. Der Frequenzbereich blieb der gleiche—aber die drehbare Welle geriet bei ihren kritischen Drehzahlen in Schwingungen, versetzte dabei das Endoskop in unerwünschte Erschütterungen, die die Aufmerksamkeit des Beobachters ablenkten, liessen auch die Drehzahl wieder schwanken und führten, wenn auch in geringem Masse, zu einer schwankenden Bewegungsfrequenz des stroboskopischen Bildes. Das Gewicht dieses Versuchsapparats einschliesslich der biegsamen Welle betrug nur 800 Gramm.

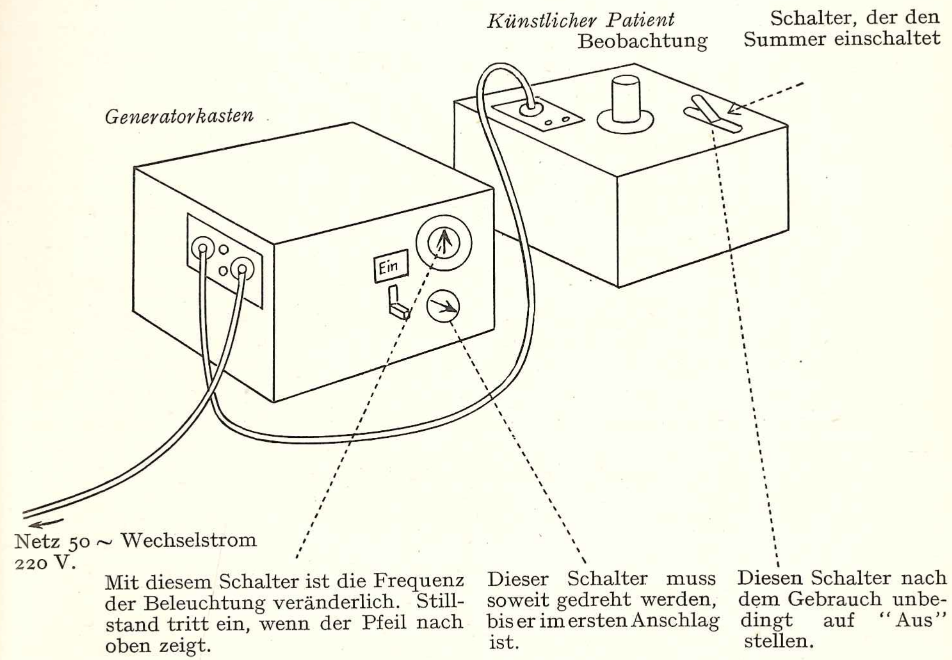

Bild 1. Skizze zur Vorführung des endostroboskopischen Effekts am künstlichen Patienten.

Inzwischen war es möglich geworden kleine Nebenschlussmotoren zu bauen. Es wurde in Folge dessen angestrebt in einer dritten Versuchsreihe das erste Modell zu verbessern, indem ein solcher Motor von nur 60 mm Durchmesser und 55 mm Länge eingebaut wurde. Der Regulirwiderstand wurde auf einem Tischchen neben dem Patienten angebracht, die Beleuchtungslämpchen wurden durch einen Schalter vom Handgriff aus betätigt. Das Gewicht war so auf 450 Gramm herabgesetzt. Aber die Hoffnung mit diesem Modell *aller* Nachteile der früheren vollends Herr zu werden, hat sich doch nicht restlos erfüllt: die Schwankungen waren geringer geworden, aber sie liessen sich nicht ganz vermeiden. So entstand der Wunsch