und das Bestreben, den Motorantrieb der Unterbrecherscheibe ganz zu beseitigen und den stroboskopischen Effekt auf anderem Wege zu erreichen.

Diese Aufgabe ist gelöst. Es ist in der Tat gelungen den stroboskopischen Effekt ohne störende Geräusche und ohne Erschütterungen hervorzurufen. Das Mittel dazu ist die Verwendung modulierten Lichtes d. h. eines Lichtes, dessen Intensität im Takte der Modulation schwankt. Die Frequenz der Helligkeitsschwankungen muss zu der Frequenz des Tones, den die Versuchsperson singt, in der bekannten Beziehung stehen, die der stroboskopische Effekt voraussetzt. Ein solches "moduliertes Licht" lässt sich beispielsweise vermittels wechselstromgespeister Glimmlampen erzeugen. Legt man an eine solche Glimmlampe etwa eine Wechselspannung mit der Frequenz 50 Hertz, so schwankt deren Lichtstärke 100 mal in der Sekunde zwischen 0 und den Grösstwerten auf und ab. Beleuchtet man mit diesem Licht einen periodischen Vorgang etwa mit der Frequenz 98, so erhält man eine Bewegungsfrequenz des stroboskopischen Bildes von 2 Hertz.

Das ist der Grundgedanke unserer neuen endostroboskopischen Methode. Indess waren noch allerlei technische Schwierigkeiten vorhanden. Zunächst mussten genügend kleine Glimmlampen hergestellt und mit einer besonderen Elektrodenanordnung versehen werden. Auch das ist wohlgelungen. Die kleinen Lampen werden mit einem Wechselstrom gespeist, dessen Frequenz im Bereich von 100–1200 Hertz in einfacher Weise verändert werden kann. Die Erzeugung dieses Wechselstroms erfolgte zuerst in einer kleinen Maschine; jetzt geschieht sie in einer Elektrodenröhrenanordnung, deren Betrieb vom Netz mit 220 Volt Spannung möglich ist: die Frequenzkonstanz des Röhrengenerators ist viel grösser als die einer Lochscheibe mit Motorantrieb.

An dem so ausgeführten Apparat sind sonst bewegte Teile irgend einer Art nicht vorhanden. Erschütterungen fallen ganz weg. Sein Gewicht ist kaum grösser als das Gewicht eines normalen Endoskops (180 Gramm); er ist also sehr leicht und bequem zu handhaben.

Soll z. B. mit dem Gerät bei einem Ton von 300 Hertz Frequenz gearbeitet werden, so wird der Ton an dem "Generator" eingestellt. Ein gewöhnlicher Telefonhörer, der neben dem Wechselstromerzeuger angeordnet ist, gibt dem Patienten den Ton an. Nach Ausschaltung des Hörers wird die Frequenz dann so weit geändert, bis das stroboskopische Bild die gewünschte Frequenz hat (1–5 Hertz).

Vor dem Vortrage wurde auf Wunsch der Kongressleitung an zwei Tagen eine Vorführung der Methode abgehalten. Hierzu bediente ich mich des Generators im Original. Statt der Versuchsperson war eine Vorrichtung verwendet, die als der "künstliche Patient" bezeichnet wurde; sie enthält eine schwingende Zunge als Nachbildung der Stimmlippen. Die in dem Kasten eingebaute Glimmlampe, die das modulierte Licht erzeugt, ist identisch mit der im Endoskop unterzubringenden Glimmlampe. Wie die beigegebene Skizze zeigt, konnte jeder der zahlreichen Teilnehmer den endostroboskopischen Effekt selbst einstellen, beobachten und dauernd kontrollieren.

47. Prof. E. W. Scripture (Vienna): *The nature of speech.*

Microphonic Speech

In Fig. 1 we have before us the track of a word on a sound film. We know nothing about it except that longitudinal distance means time and transverse movement of the upper edge means something doing in speech. We proceed to study it as if we had been born deaf and had never read anything about speech, sound or hearing.

The first line shows no change in the upper edge of the track; it is the registration of a time of *nothing doing*. After the moment indicated by the small vertical arrow the track shows up and down movements. These movements we term *vibrations*. As the apparatus was constructed to record with great magnification, the original vibrations must have been microscopic in size; we therefore give the term *microphonic speech* to what was recorded.

After the vertical arrow the second line begins with sharp and rapid vibrations with a mixture of regularity and irregularity; the vibrations start strong and then become weaker. Such a group of vibrations we term an *explosive stretch*. Measurements give the numbers for the duration of the stretch and for the heights and times of occurrence of the vibrations. After the moment marked by the horizontal arrow there follows a rather long stretch of weak vibrations showing a mixture of regularity and irregularity; this we term a *rough stretch*. Measurements give the numbers for the duration of the stretch and the heights and times of recurrence of the vibrations.

After the point marked ~ there occurs a brief rise and fall of the zero level of the vibrations. This we term a *zero waver of the rough stretch*.

After the point marked 1 the line of vibration shoots up suddenly and then descends to zero in a series of vibrations. This is repeated more strongly after 2, again more strongly after 3, and so on. Such a movement we term a *vibration bit*. The zero waver is found in each vibration bit. When the vibrations are sharp and strong we speak of a *lively movement*. When they are smooth as in the sixth line we speak of a *billowy movement*.

The small vibrations in most of the seventh line entitle us to speak of a stretch of *almost nothing doing*. The line ends with a weak *explosive stretch* and a *rough stretch*. This persists in a very faint form to the end of the record. Measurements give the numbers that characterize these stretches.

The tracks of a number of vowels spoken by the person who made the record in Fig. 1 are reproduced in Fig. 2. Comparison of the tracks establishes the following principles:

1. A vowel stretch consists of a series of contiguous vibration bits.
2. The time-profiles of the vibration bits in a vowel stretch change continuously, gradually and progressively.
3. The vibration bits in different vowel stretches have different time-profiles.
4. Every vibration bit begins strong and fades rapidly to zero.

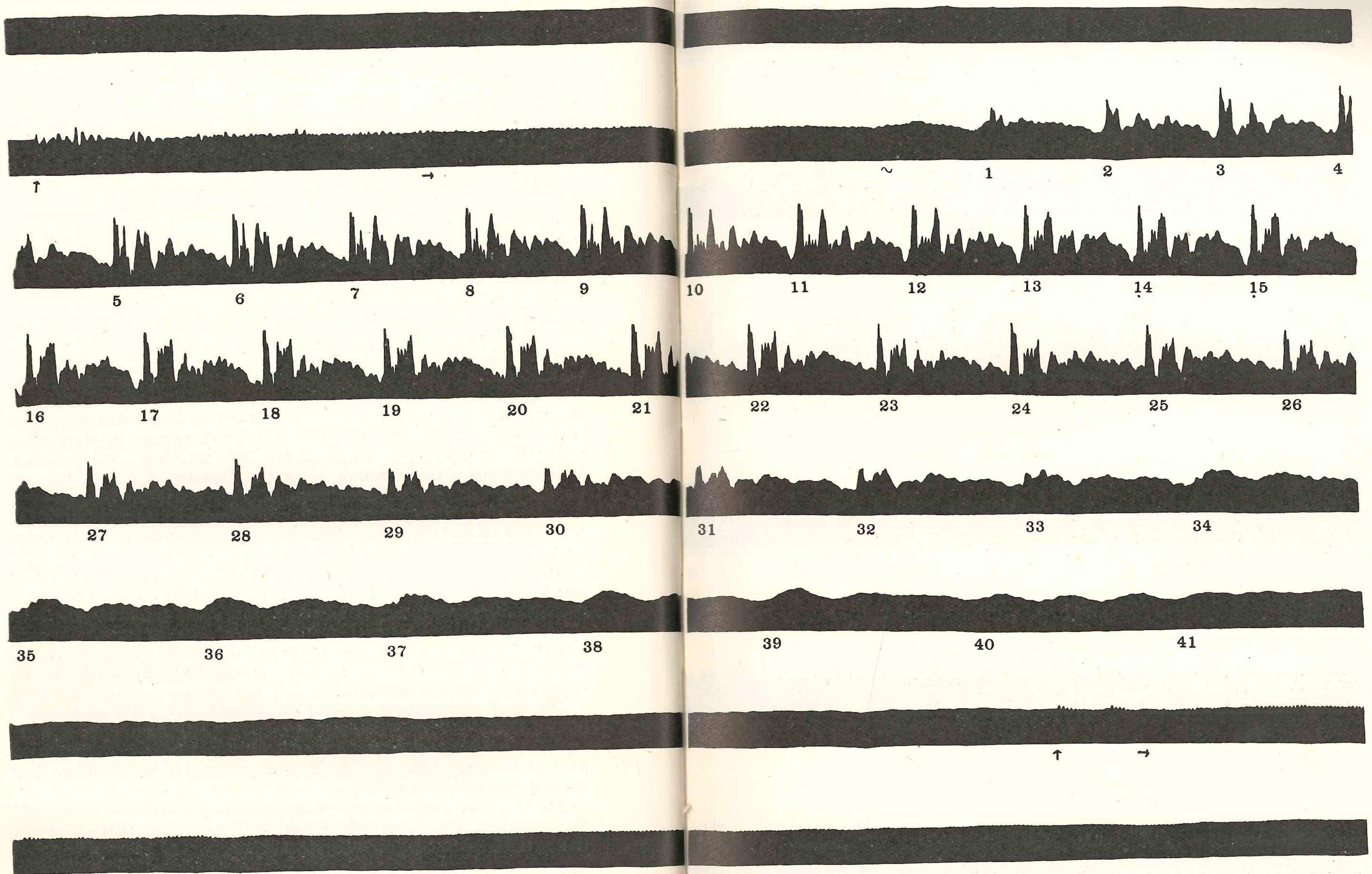

Fig. 1. Sound track of *cant*.

The eye recognizes the profiles of the different vowel stretches just as they stand; it does not analyse them any more than it analyses the portraits of persons. At the most we can make only some general statements such as that the profiles for *oo* look "smooth", those for *aw* and *oh* "toothy", *ah* and *ay* "scratchy" and those for *ee* "furry". A person learns readily to recognize the different profiles; if he would devote time to the matter he could learn to read a speech track just as he reads a book.

By measuring the heights of the peaks in a profile, and their times of occurrence and also the times of recurrence of smoother movements of vibration, we get the numbers that characterize the profiles. This method of *peak and centroid analysis* rests upon no hypothesis; it is simply a method of expressing the profiles in numbers.

Our knowledge of microphonic speech is advancing rapidly. We used to think that vowels were characterized by the presence of certain frequencies or groups of frequencies. Recent experiments have shown that any frequency or group of frequencies or range of frequencies can be removed from vowel vibrations without affecting the vowel character; as long as any recognizable sound is left, the vowel character remains the same. Now we know that the vowel character does not depend on the frequencies but on the profile of the vibration and that the profiles are perceived without analysis.

Another piece of knowledge concerns the nature of a speech stretch. We used to think that it consisted of a series of little blocks, termed "speech sounds", united by little slides of speech cement, termed "glides". We now know that the stretch of speech between two pauses is continuous with no breaks of any kind. We follow with our eyes the change of the character of a speech sound and find that it passes gradually into the character of the following speech sound; we find that speech does not consist of blocks of constant sounds but of a continuous sound that changes more or less gradually from beginning to end.

Macrophonic Speech

The vibrations of microphonic speech require something in the speaker to produce them. As they come from his mouth and nose we enquire if there is anything else issuing from the same sources. For this purpose we place a funnel with a rubber aircushion around its edge closely over the mouth and a glass tip in one of the nostrils. A wide tube leads from the mouth funnel to a membrane of soft oiled silk whose movements are enlarged and registered by a light lever on a moving smoked surface. The membrane follows changes in the air pressure; it does not respond to the vibrations of microphonic speech. A small rubber tube leads from the glass tip to a similar registering apparatus. The entire system is shown in Fig. 3. The apparatus registers the mass movements of the air from mouth and nose and nothing else. These mass movements are found to constitute a complete system of speech. On account of the large dimensions of the movements this system is termed *macrophonic speech.*

In the macrophonic registrations reproduced in Fig. 4 the *nasal registration* shows at the beginning a very slight breath movement followed by a stretch of nothing doing. Thereafter follows a continued slight breath movement with small waves. After a stretch of waves slowly increasing in amplitude the breath line rises to a considerable height and then falls to zero. During the rise the waves increase strongly in amplitude; during the fall they fade away. After a stretch of zero line the breath line rises and then sinks slowly. The *mouth registration* indicates nothing doing till the point marked 3. Here an emission of breath begins, rises to a maximum and then falls. In the course of the fall waves appear. These quickly become very strong.

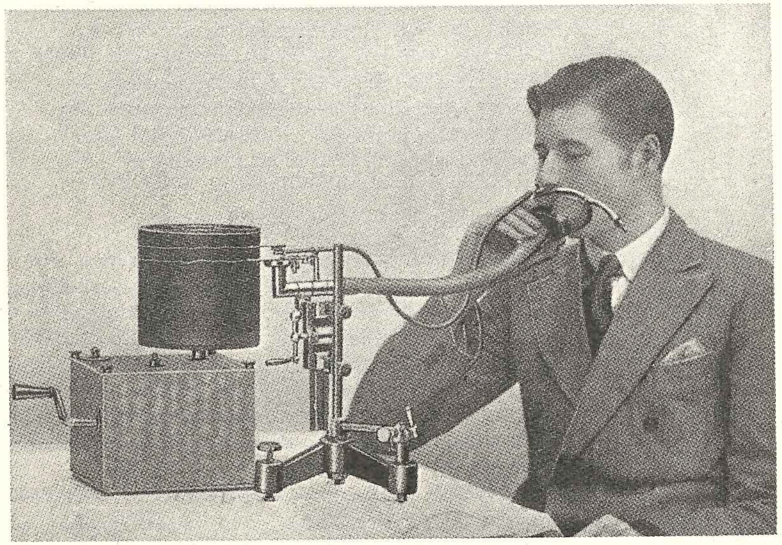

Fig. 3. Making macrophonic registrations of speech.

After the point marked 9 the waves diminish rapidly in amplitude. They are still traceable up to the point marked 12. A stretch of nothing doing from 12 to 13 is followed by a sudden strong emission of breath that rises to a maximum and fades slowly away.

The registrations show currents, sudden emissions, stoppages and waves of breath. A sudden emission may be termed appropriately an *explosion*. The waves of breath may be termed *puffs*. The waves in the registrations are not records of microphonic vibrations, as the membranes do not respond to such vibrations.

Although the speech track in Fig. 1 and the registrations in Fig. 4 were not made by the same person, there is close correspondence between the parts of each. The stretch of nothing doing in the mouth registration corresponds to the stretch of nothing doing in the speech track. The stretch from 3 to 6 corresponds to the stretches of explosion and roughness in the track. The first faint wavy movement in the mouth registration corresponds to the brief stretch indicated by ∼ in Fig. 1. The waves from 7 to 12 correspond to the vibration bits in the track. If both registrations had been made simultaneously there would be one wave for each bit. The straight line, the upward rise and the slow fall in the mouth line correspond to the stretch of almost nothing doing, the explosions and the final noise in the track.

Holding a small box close to my mouth I let a sharp puff of air escape from my lips; the result is an explosive noise. This happens because the sharp puff—or explosion—sets the air in the cavity into vibration. With larger or smaller boxes the explosive noise alters its character and we conclude that the profile of the vibrations depends on the size of the cavity. A steady current from the lips produces a maintained noise from the box; the character of the noise depends on the size of the box. We can readily understand that a sharp release of breath at some point in the mouth passage will produce an explosive noise and a steady current of sufficient strength will produce a maintained noise; in both cases the character of the noise will be determined by the size, form, openings and walls of the vocal cavity.

By snapping the thumb out of a flask we can produce a brief rapidly fading tone. By snapping it out of the mouth we produce a momentary vowel. The snap in each case produces a sharp rarefaction that sets the air in the cavity into vibration. We can produce a momentary vowel by closing the glottis and snapping it open. By rapidly repeating this action we produce the vowels of speech.

The currents, sudden emissions and puffs of breath of macrophonic speech arouse the air of the vocal cavity to vibration. The profiles of the vibrations depend on the size, shape, openings and walls of the cavity.

A speech sound appears in a series of forms: the microphonic sound, the macrophonic activity and the myokinetic activity. We can go back through the neurokinetic activity and ultimately arrive at the inner activity.

The muscular activity that produces the air current that constitutes macrophonic speech and regulates the shape, size, openings and walls of the vocal cavity in microphonic speech may be termed *myokinetic speech*. The relations among the three systems may be stated as follows. The profiles of microphonic speech with the two dimensions of time and something vibratory doing are produced by the impulses of macrophonic speech with the three dimensions of time, mouth breath and nasal breath. These in turn are produced by myokinetic speech with the dimension of time and many dimensions of muscular contraction.

The Perception of Speech

When a wheel with a series of openings is rotated in front of a jet of air, a brief noise is heard each time an opening passes the jet. Physically the experiment consists in producing a series of puffs of air rubbing over an edge. Mentally it consists in perceiving a series of brief noises. One thing is common to the two series, namely, the intervals at which the puffs and noises occur. The number of times an interval could be repeated in 1 second is termed the frequency. The common thing in the two series may be expressed by the frequency number. As the disc is speeded up, the brief noises occur more frequently. Above a certain frequency a *tone* is heard. In the physical series nothing is added and nothing is changed except the frequency number. To the mental series something entirely new is added, namely, the sensation of tone. The tone is a purely mental affair to which nothing except the frequency number corresponds in the physical series. In speech this is termed the *voice tone* because it corresponds to the frequency of occurrence of the vibration bits in microphonic speech and of the puffs in macrophonic speech.

The vibrations of microphonic speech are perceived as *speech characters* and *speech qualities*; there is never any analysis. Vowels are perceived as alike or different according to their profiles. Speech-sounds of different vibratory forms are perceived with resulting sensation differences. Some vibration profiles produce tonelike sensations because their peaks and centroids have frequency numbers.

The Overlapping of Speech Sounds

The person who spoke what appears as the track in Fig. 1 wrote *cant* on the piece of film. Tracks from the same speaker labelled by him *can*, *cat*, etc. begin with stretches of nothing doing, explosion and noise similar to those in Fig. 1. Tracks labelled *ash*, *an*, etc. begin with vibration stretches like the first part of the stretch with 41 vibration bits in Fig. 1. Tracks for *an* and *can* ended with vibration stretches like the latter part of the stretch with 41 bits. The track for *cat* ended in a manner similar to that in Fig. 1. These comparisons lead to the conclusion that the impulse of the speaker in producing the track in Fig. 1 was the sum of four different impulses. Let us label these impulses **k**, **æ**, **n**, **t**.

Part of the condition of nothing doing registered in the first line of Fig. 1 and the vibratory condition from the beginning of the second line up to the point marked 1 belong to the expression of the first of the speaker's impulses, namely, to **k**. As the small mixed irregular-regular vibrations can be followed through most of the vibration bit 1 this bit must correspond to the last part of **k**. The series of vibration bits from the point marked 1 to beyond the middle of the fifth line belong to the expression of the inner impulse **æ**. The smooth vibration bits that appear clearly after the point marked 34 belong to the expression of **n**. The smooth character can be traced backward through the preceding six bits. The bits from 1 to 34 belong evidently to the expression of **æ** and those from 27 to 41 to the expression of **n**. The condition of almost nothing doing that occurs in most of the seventh line is already anticipated in the last few vibratory bits assigned to **n**. The double explosion belongs to **t**. The noise with which the registration ends is so long that it cannot be treated merely as the finish of the explosion of **t**. It is, in fact, the registration of a final noise like an **h** or an aspirate with which the word closes. The track in Fig. 1 is thus the expression of the five sounds **k**, **æ**, **n**, **t**, ^{c}. They do not, however, occur in succession but are overlapped.

The speaker, whose macrophonic registrations are reproduced in Fig. 4, indicated what he had intended to say by writing *cant* on the registration sheet. Comparisons of various registrations by

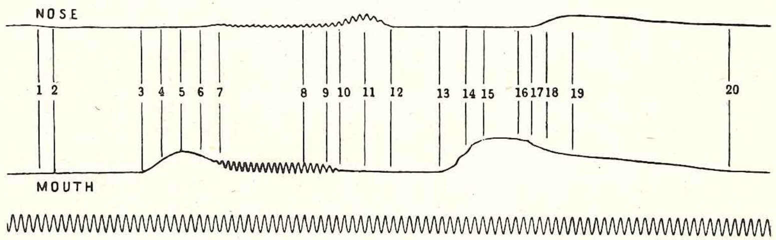

Fig. 4. Macrophonic registrations of *cant*.

this speaker showed that they contained portions resembling one another that he labelled *k* or *c* or *ck*. These all agreed in showing a mouth stoppage ending in a fairly sharp explosion followed by a more or less rapidly decreasing breath current. The impulse that he expressed in these cases we designate by **k**. In Fig. 4 the decrease of breath can be traced to the point marked 8. The waves from 7 to 10 register the puffs of a vowel sound. From the fact that the speaker wrote the word *cant* we infer that it was the vowel used in *as*, *can*, etc.; this we indicate by **æ**. The waves for **æ** begin high up on the breath line that finishes **k**. This indicates that the first part of **æ** overlapped the latter part of **k**. The waves in the mouth line decrease more rapidly after 9 than they would in a vowel alone; this indicates that the sound **n** began before the vowel was finished. The **n** extends

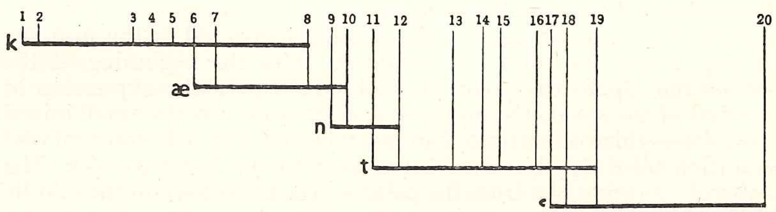

Fig. 5. Overlapping of speech sounds in Fig. 4.

until 12. The great faintness of the last part of **n** in the mouth registration and the fall of the breath level in the nasal registration show that the stoppage for **t** was already beginning at 11. Mouth and nasal stoppage followed by a mouth explosion and a subsequent fading breath were the usual registrations of **t** by this speaker. The word ends with a long emission of breath in both mouth and nasal registrations. It may be indicated by ^{c}. The overlapping of the speech impulses in the macrophonic registrations is shown in Fig. 5.

Overlapping is an inner, or psychic, matter that happens before the impulses are expressed in speech. The impulses must occur more or less simultaneously. Inner activity, or mind, must therefore possess a dimension of simultaneity in addition to those of time and intensity. This means simply that several things can occur in the mind at the same time. Since the speech sounds constituting a word occur more or less simultaneously a word must occur not only in a stretch of time but also with an extension in the *dimension of simultaneity*. Simultaneity is for mind what a space dimension is for material things. In the expression of the internal impulses there is no dimension of simultaneity; whatever is in the mind at any one moment must be expressed as one thing. This is what gives rise to the phenomenon we have termed overlapping.

Conclusion

In conclusion I wish to give two quotations:

"When you can express what you are speaking about in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the stage of science." (Thompson, *Popular Lectures and Addresses* (London, 1899), I, 73.)

"Something is doing—we don't know what....From so unpromising a beginning we really do get somewhere. We bring into order a host of apparently unrelated phenomena; we make predictions, and our predictions come off. The reason—the sole reason—for this progress is that our description is not limited to unknown agents executing unknown activities, but numbers are scattered freely in the description....Out of the numbers proceeds that harmony of natural law which it is the aim of science to disclose." (Eddington, *Nature of the Physical World* (Cambridge, 1928), p. 291.)

The first of these quotations may be restated in the following way: When you meet a phonetic fact, measure it and say, "I've got your numbers". The other one says to the phonetic fact, "Call yourself what you please; it makes no difference; I've got your numbers".

Measurements of one kind give us the numbers that constitute microphonic speech; measurements of another kind give us the numbers for macrophonic speech. Still other numbers give us myokinetic speech, that is, the muscular activity involved in speaking. A special set of numbers constitutes neurokinetic speech, that is, the action of the nervous system in speech. For inner speech there are still other numbers. There are also special systems of numbers for linguistics, metrics and all sciences that have to do with speech. Phonetics is the science that includes all the numbers constituting speech.

48. Prof. P. Menzerath (Bonn): *Neue Untersuchungen zur Steuerung und Koartikulation.*

In meiner Arbeit über *Steuerung und Koartikulation* (Bonn, Dümmler, 1933) konnte ich mit Dr. de Lacerda die Unzuverlässigkeit in der artikulatorischen Ausdeutung der Mundstromkurven, der Labiogramme usw. nachweisen. Zum anderen ersetzten wir auf Grund unserer Ergebnisse die alte Dreiphasentheorie der Lautartikulation durch eine Zweiphasentheorie. Wir leugneten die frühere Auffassung, die Anglitt-Stellung-Abglitt schied, und setzten an ihre Stelle die Zweiphasentheorie. So ist die konsonantische Artikulation stets eine Schliessungs-Öffnungsbewegung, mit einem artikulatorischen Maximum zwischen diesen beiden Phasen; die vokalische ist eine Öffnungs-Schliessungsbewegung, mit einem artikulatorischen Minimum in der "Fuge".

Alle Artikulation aber verläuft in Dauerbewegung und wiederum so, dass die Art der Bewegung vom folgenden Laut bestimmt wird und die Bewegungs*folge* durch die *struktur*gemässe Artikulationsverbindung, die ich als Artikulations*verflechtung* bezeichnete und womit ich sagen wollte, dass das alte Bild der "Kette", der "Anreihung", in der die Lautfolge erscheinen soll, nicht zutrifft und durch ein anderes zutreffenderes Bild, z. B. des "Geflechtes" ersetzt werden muss. Also noch einmal: die Silbe (das Wort) ist, artikulatorisch gesehen, eine stufenweise realisierte Gesamtheit, eine Ganzheit, deren Teile sich in jedem Moment anders verbinden im Hinblick auf das erstrebte Ziel, die Verbindung von Lauten zur Silbe.

Seit der Zeit der Veröffentlichung—für die mir keine irgendwie massgebende oder stichhaltige Ablehnung bekannt geworden ist—habe ich reichlich Versuche gemacht, um diese Hypothese nachzuprüfen und sie, wenn nötig, durch eine neue zu ergänzen.

Damals gewannen wir—Einzelheiten darf ich wohl als bekannt voraussetzen—unser kritisches Material mit Hilfe von Labiographen, die entweder Dauerschreibung zeigten (um die Bewegungsfolge zu registrieren), oder Momentschreibung (Elektrolabiogramm; zur Entscheidung über den berühmten Verschlusspunkt und den nicht minder berühmten "Kurvenknick"). Zur Zeit bevorzuge ich, der einfachen Registrierung halber, und der leichter zu handhabenden Apparatur wegen, das *Geniogramm* (die Bewegungen des Kiefers sind allerdings genau so zusammengesetzt wie bei den Lippen: auf-ab, vor-zurück; das Labiogramm wie das Geniogramm geben nur die Auf-Abbewegung). Eine Reihe von Kurven habe ich ausgestellt. Ich will Ihnen auch ein paar im Lichtbild zeigen. Sie können hier feststellen, dass die sogenannten charakteristischen Punkte der Mundkurve auf dem Geniogramm nur in seltenen Fällen eine Entsprechung haben. Keinesfalls aber ist die Mundkurve von *sich* aus *unmittelbar* artikulatorisch zu deuten. Ihre artikulatorische Interpretation erfordert stets eine Reihe von Parallelkurven, von denen man im Interesse der Sicherheit nicht abgehen sollte.

Doch scheint mir, dass unsere Zeit auch von der Phonetik feinere Untersuchungsmethoden verlangt. Wir können eigentlich nicht mehr

mit ruhigem Gewissen bei unserem heissgeliebten Kymographen stehen bleiben, obschon ich weit davon entfernt bin, diese Apparatur als vollkommen nutzlos abzulehnen. *Das Problem bestimmt stets die Methode*; mit Kanonen schiesst man nicht nach Spatzen.

Ich möchte Ihnen einige solche Versuche mit besseren Methoden vorführen und nehme die von anderen im gleichen Sinne gemachten Anstrengungen als gegeben hin, ohne sie im einzelnen zu bezeichnen. Im wesentlichen arbeiten wir heute am aussichtsreichsten mit dem Tonfilm und der Röntgenphotographie (im weitesten Sinne).

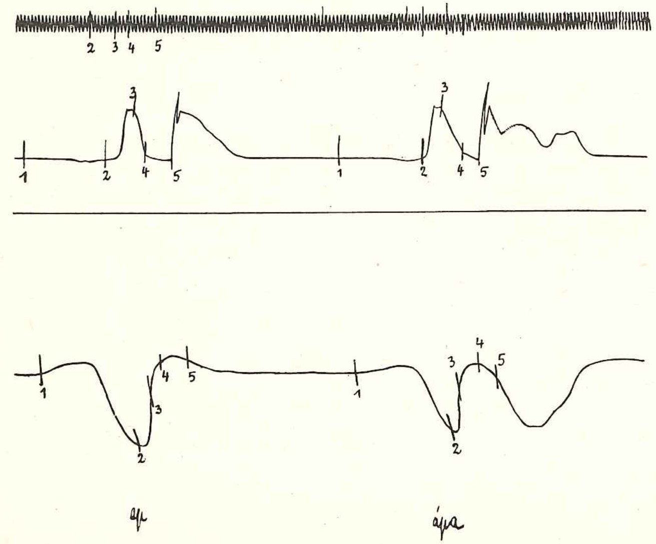

Bild 1.[1] **ap** **ápa**.

ap Die auf beiden Kurven entsprechenden Punkte sind jeweils mit den gleichen Zahlen versehen. **a** in **ap** beginnt bei 2, endet vielleicht bei 3, das aber—nach Ausweis des Geniogramms—hier keinesfalls der Verschlusspunkt ist. 4 und 5 gehören zum Verschluss, dessen Maximum genau in der Mitte liegt. Die Bewegung im Verschluss ist unten sehr klar zu sehen.

ápa Bei **ápa** wird das Bild noch komplizierter. Bei 3 ist der Verschluss noch nicht erreicht, 4 ist das Maximum; der Beginn des Verschlusses ist also gar nicht feststellbar. Bei 5 setzt das Schluss-**a** ein.

(1) *Röntgenstereoskopie.* Dem Phonetiker, dessen Auge i. allg. an die Deutung von Röntgenbildern nicht gewöhnt ist, erscheint eine Röntgenaufnahme zunächst nur flächenhaft, nicht plastisch. Diesen Nachteil habe ich durch die Röntgenstereoskopie zu beheben getrachtet, von der Ihnen in der Ausstellung ein paar gut gelungene Proben zugänglich sind. Wir kommen gleich in anderem Zusammenhang noch darauf zurück. Das gewöhnliche Röntgenbild zeigt uns aber—wie auch das Palatogramm—nur Stellungen, nicht aber Bewegungen. *Die* aber wollen und müssen wir haben, und zwar,

[1] Die Zeitmarkierung (1/50 sec.) ist bei den 4 Abbildungen dieselbe.