

Zeitschrift:	Technische Mitteilungen / Schweizerische Post-, Telefon- und Telegrafenbetriebe = Bulletin technique / Entreprise des postes, téléphones et télégraphes suisses = Bollettino tecnico / Azienda delle poste, dei telefoni e dei telegrafi svizzeri
Herausgeber:	Schweizerische Post-, Telefon- und Telegrafenbetriebe
Band:	39 (1961)
Heft:	5
Artikel:	Phonétographe: présent et futur : niveleur, formants-suiveurs, sons et couleurs, automation et phonaction
Autor:	Dreyfus-Graf, J.
DOI:	https://doi.org/10.5169/seals-875247

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

Download PDF: 24.01.2026

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Jede Stelle kann über das Telexnetz jeden Teilnehmer des In- und Auslandes erreichen. Im Normalzustand ist die Hauptstelle mit der Amtsleitung verbunden. Wünscht die Nebenstelle eine Verbindung aufzubauen oder muss ein Anruf von einem Telexteilnehmer zur Nebenstelle weitergeleitet werden, wird die Amtsleitung in der Hauptstelle zur Nebenstelle durchgeschaltet. Die Durchschaltung bleibt bestehen, bis die Nebenstelle oder der rufende Telexteilnehmer durch das Schlusszeichen die Verbindung auslöst. Will die Hauptstelle während einer Verbindung Nebenstelle-Telexteilnehmer eine Verbindung aufzubauen, wird die Umschaltung erst wirksam, wenn die erwähnte Verbindung abgebaut ist. Dasselbe gilt, wenn die Nebenstelle während einer Verbindung Hauptstelle-Telexteilnehmer umzuschalten wünscht. Der Fernschreibstellen-Umschalter leistet jedem Geschäftsmann wertvolle Dienste, der von zwei Orten aus Meldungen abzusenden und zu empfangen wünscht, so etwa dem Börsenmakler, der sich abwechselnd in seinem Bureau und an der Börse befindet.

Abschliessend sei erwähnt, dass die Fernschreib-

zentrale 4/10 in etwas grösserem Ausbau schon seit mehreren Jahren bei den Siemens-Dienststellen Berlin, München und Erlangen mit bestem Erfolg eingesetzt ist. Die Fernschreibzentrale 5/5 wurde an der letztjährigen Messe in Hannover erstmals der Öffentlichkeit vorgeführt. Derartige Anlagen erlauben, den Fernschreibverkehr weiter zu rationalisieren und die Anwendungsmöglichkeiten für Fernschreiber zu erweitern.

Bibliographie

- [1] *Miemiec, Horst.* Das Fernschreibnetz der Siemens-Werke. Siemens-Zeitschrift Nr. 12/1958, S. 826ff.
- [2] *Machutt, Heinz-Joachim* und *Miemiec, Heinz.* Neue Möglichkeiten für den Anschluss von Telex-Teilnehmern mit mehreren Fernschreibern. Siemens-Zeitschrift Nr. 4/1960, S. 184ff.
- [3] Ferner folgende Drucksachen der Siemens & Halske AG., Fernschreibtechnik:
 - a) TW-Nebenstellenanlage 1/1 für Ortsleitungen, Beschreibung (Juli 1956);
 - b) Siemens-Fernschreibzentrale ZbT 5/5 beim Teilnehmer (Prospekt);
 - c) Siemens-Fernschreibzentrale ZbT 1/1, Fernschreibstellen-Umschalter ZbT U (Prospekt).

J. Dreyfus-Graf, Genève

Phonétographe: Présent et futur *

681.841.3

(Niveleur, formants-suiveurs, sons et couleurs, automation et phonaction)

Résumé. Le terme «sonographe» désigne diverses machines (brevets publiés ou déposés) capables de transformer des suites de sons en des successions de graphiques analogiques ou numériques, par des oscilloscopes ou des imprimantes.

En particulier, le sonographe phonétique (ou phonétographe) transforme les langages parlés en langages écrits à l'aide de signes alphabétiques usuels.

La partie «analyseur» des sonographies imite certains caractères du récepteur auditif humain. Elle discerne cinq catégories d'éléments d'information: l'intonation, l'harmonie, la mélodie, le bruit et le rythme.

Cela à l'aide des organes suivants: compresseur sélectif ou nivelleur sonémique, filtres de formants quasi fixes, de formants mobiles ou suiveurs, et de sub-formants.

Les éléments d'information (binaires, ternaires ou doubles binaires) sélectionnés par ces organes sont au nombre de 18, par cinquième de seconde, dans le phonétographe, prototype III. Ils seront une cinquantaine, par quinzième de seconde, dans le prototype IV.

La «serrure» est un modèle mécanique illustrant le fonctionnement de l'analyse combinatoire, de la théorie de l'information, et de la partie numérique des sonographies, entre autres.

Le phonétographe III permet d'écrire des mots sous dictée, à condition que le parleur adapte sa diction à la machine, c'est-à-dire qu'il contrôle sa hauteur de voix, son intensité et son rythme.

Le phonétographe IV, au contraire, sera pourvu d'automatismes qui adapteront la machine à divers parleurs ou parleuses.

Parmi les applications pratiques des phonétographes, on peut citer, à titre d'exemples: machines de bureau, téléphonie sur bande étroite, commandes orales de téléimprimeurs, de calculatrices, de traductrices, de trieuses et de toute programmation. D'autre part, ils servent à des analyses acoustiques de toutes sortes.

Zusammenfassung. Der Ausdruck «Sonograph» bezeichnet verschiedene Maschinen (Patente veröffentlicht oder angemeldet), die Tonfolgen mit Hilfe von Oszillographen oder Schreibmaschinen in Zeichenfolgen umwandeln. Insbesondere verwandelt der phonetische Sonograph (oder Phonetograph) Wortsprache in Schriftsprache, unter Verwendung der üblichen alphabetischen Zeichen.

Der Analysator-Teil des Phonetographen ahmt gewisse Funktionen des menschlichen Hörapparates nach. Er unterscheidet fünf Nachrichtengruppen – Betonung, Harmonie, Melodie, Geräusch und Rhythmus – mit Hilfe folgender Organe: selektiver Dynamik-Presser (oder sonemischer Kompensator), Filter für quasi-feste, bewegliche oder nachfolgende Formanten sowie für Subformanten.

Beim Prototyp III des Phonetographen beläuft sich die Zahl der binären, ternären und doppelbinären Nachrichtenelemente auf 18 in der Fünftelsekunde; beim Prototyp IV werden es etwa 50 in der Fünfzehntelsekunde sein.

Das «Schloss» ist unter anderem ein mechanisches Modell der Kombinatorik, der Nachrichten-Theorie sowie des numerischen Teils des Sonographen.

Der Phonetograph III schreibt Worte nach Diktat, jedoch muss sich der Sprecher der Maschine anpassen, das heißt, er muss Stimmlage, Intensität und Rhythmus kontrollieren. Hingegen wird der Phonetograph IV verschiedene automatische Hilfsorgane enthalten, welche die Maschinen an jeden Sprecher oder jede Sprecherin anpassen werden.

Unter den praktischen Anwendungen der Phonetographen sollen nur die folgenden erwähnt werden: Büromaschinen, Schmalband-Telephonie, mündliche Betätigung von Telegraphen, Rechen-, Übersetzungs- und Sortiermaschinen, ferner Klang-Analysatoren jeder Art.

* Compte rendu des conférences ayant accompagné les premières présentations du phonétographe, prototype III, à l'Institut de physique de l'Université de Genève, le 9 décembre 1960 (Sonographe et serrure) et le 24 mars 1961 (Phonetographe: Présent et futur).

A la base de tout progrès, il y a l'information. Elle met en regard ce qui existe et ce qui est possible. Elle permet de faire un choix. La cybernétique a démontré que la mission primordiale de l'information résidait dans la mesure de l'erreur, afin d'en permettre la correction.

L'information humaine repose essentiellement sur le langage parlé et sur sa mise en mémoire lisible, le langage écrit dont le type le plus puissant est alphabétique. En effet, celui-ci se contente d'une trentaine de signes graphiques très simples pour afficher ou conserver les myriades de pensées, vraies ou fausses, de l'humanité présente ou passée. L'invention de l'alphabet – qui était purement phonétique à l'origine – date de l'an 1500 avant Jésus-Christ environ, d'une époque qu'on peut supposer contemporaine des Tables du Sinaï.

La transformation automatique de tous les sons du langage parlé en textes alphabétiques s'est heurtée, jusqu'à présent, à l'une de ces murailles qui semblaient infranchissables.

1. Le récepteur auditif

La physique est la science des phénomènes mesurables. Physiquement, un phénomène n'existe pas sans un récepteur susceptible de le mesurer. Or, l'ultime récepteur qui réside toujours dans les sens et le cerveau humain est très imprécis en raison même de son universalité. Il a besoin de prolongements spécialisés, tels que les microscopes, les spectromètres, les oscilloscopes, les calculatrices et toute la multitude des autres «scopes», «graphes», «mètres» et «trices», pour découvrir et se concilier les lois de la nature.

Le domaine des sons appartient, en dernière instance, au récepteur auditif, composé de l'oreille et du cerveau. Sa puissance d'analyse dépasse l'entendement de son propriétaire lui-même. Mais son fonctionnement n'est pas linéaire et ses mesures sont souvent fantaisistes.

Les mathématiciens, tels que *Fourier*, nous enseignent que chaque son – si complexe soit-il – peut être considéré comme la somme d'une infinité de sons purs. Malheureusement, les sons purs n'existent pas plus dans la nature que l'infini, et dès qu'il s'agit d'enchaînement de sons complexes, tels que dans le langage parlé, les mathématiciens sont incapables de nous dire où un son commence et où il finit, ni quel choix il faut faire parmi l'imbroglio de composantes théoriques.

C'est pourquoi nous sommes contraints d'inventer des récepteurs de mesure spécialisés, imitant en plus précis certains caractères du récepteur humain, pour faire progresser l'identification des sons et leur physique en général.

Résumons, pour commencer, selon le diagramme de la *figure 1*, les facultés du récepteur auditif:

Concernant les fréquences perçues, son domaine s'étend sur 10 octaves, depuis les 16 c/s du tuyau d'orgue le plus grave jusqu'aux 16 000 c/s du sifflement le plus aigu. Le piano se borne à 7 octaves et

demie, de 27 à 4096 c/s. Concernant les amplitudes vibratoires associées à ces fréquences perçues, elles peuvent varier dans le rapport de 1 à 100 000 en moyenne, entre le pianissimo d'un murmure et le fortissimo d'un clairon.

L'immensité de ces domaines perceptibles comme sons résulte du fonctionnement logarithmique du récepteur auditif: il ne ressent pas comme intervalles égaux des différences égales entre deux fréquences ou entre deux amplitudes, mais bien des rapports égaux. Par exemple, il a l'impression d'avoir parcouru le même chemin en passant de 100 à 200 c/s, dont la différence est 100, qu'en allant de 1000 à 2000 c/s dont la différence est 1000, comme s'il ne s'intéressait qu'au rapport 200:100, puis 2000:1000 qui est toujours 2. Le plus petit intervalle perceptible s'appelle le «savart» pour les fréquences, et «le décibel» pour les amplitudes. Le «savart» est 25 fois plus fin que le demi-ton du clavier tempéré des musiciens, car il divise l'octave en 300 intervalles au lieu des 12, qui se retrouvent d'ailleurs dans la musique moderne dodécaphonique ou atonale. Les inexactitudes tolérées, par exemple, dans l'accordage d'un piano sont de 3 savarts. Quant au «décibel», il divise en 100 niveaux perceptibles le domaine des amplitudes, qui peuvent varier objectivement dans les rapports de 1 à 100 000, en moyenne.

En moyenne, le récepteur auditif peut discerner 3000 hauteurs de sons ou tons et, dans chacun d'eux, 100 niveaux d'amplitude vibratoire. On peut le comparer à un piano passif possédant 3000 cordes marquées de 0 à 3000 savarts, et dont chacune peut vibrer par résonance jusqu'à une amplitude dont le niveau est gradué entre 0 et 100 décibels, en moyenne. On peut aussi l'assimiler à un peigne comprenant 3000 dents ou lames vibrantes. Le spectre auditif indique à chaque instant les cordes ou dents qui entrent en résonance et leurs amplitudes relatives.

Bien que le nombre infini des composants mathématiques, selon *Fourier*, soit déjà réduit au nombre de 3000, il serait, cependant, malaisé de construire un spectrographe donnant les mêmes informations que l'appareil auditif. Il faudrait construire 3000 résonateurs mécaniques ou filtres électriques et en faire inscrire les amplitudes en fonction du temps par un oscillographe à 3000 voies qui serait donc 500 fois plus large que celui d'un modèle courant à 6 voies. Son papier enregistreur aurait la largeur d'environ 300 mètres.

Pour en revenir au récepteur auditif naturel, sa capacité d'information par quinzième de seconde est obtenue en calculant toutes les combinaisons possibles des 3000 cordes et de leurs 100 niveaux. L'analyse combinatoire enseigne que cette capacité, c'est-à-dire le nombre total de sons discernables *S*, est le nombre 100 multiplié 3000 fois par lui-même, ce qui revient à multiplier 6000 fois par lui-même le nombre 10.

On peut se poser les questions théoriques suivantes: Si le nombre des niveaux discernables était réduit de 100 à 3, voire à 2 par corde, combien faudrait-il de

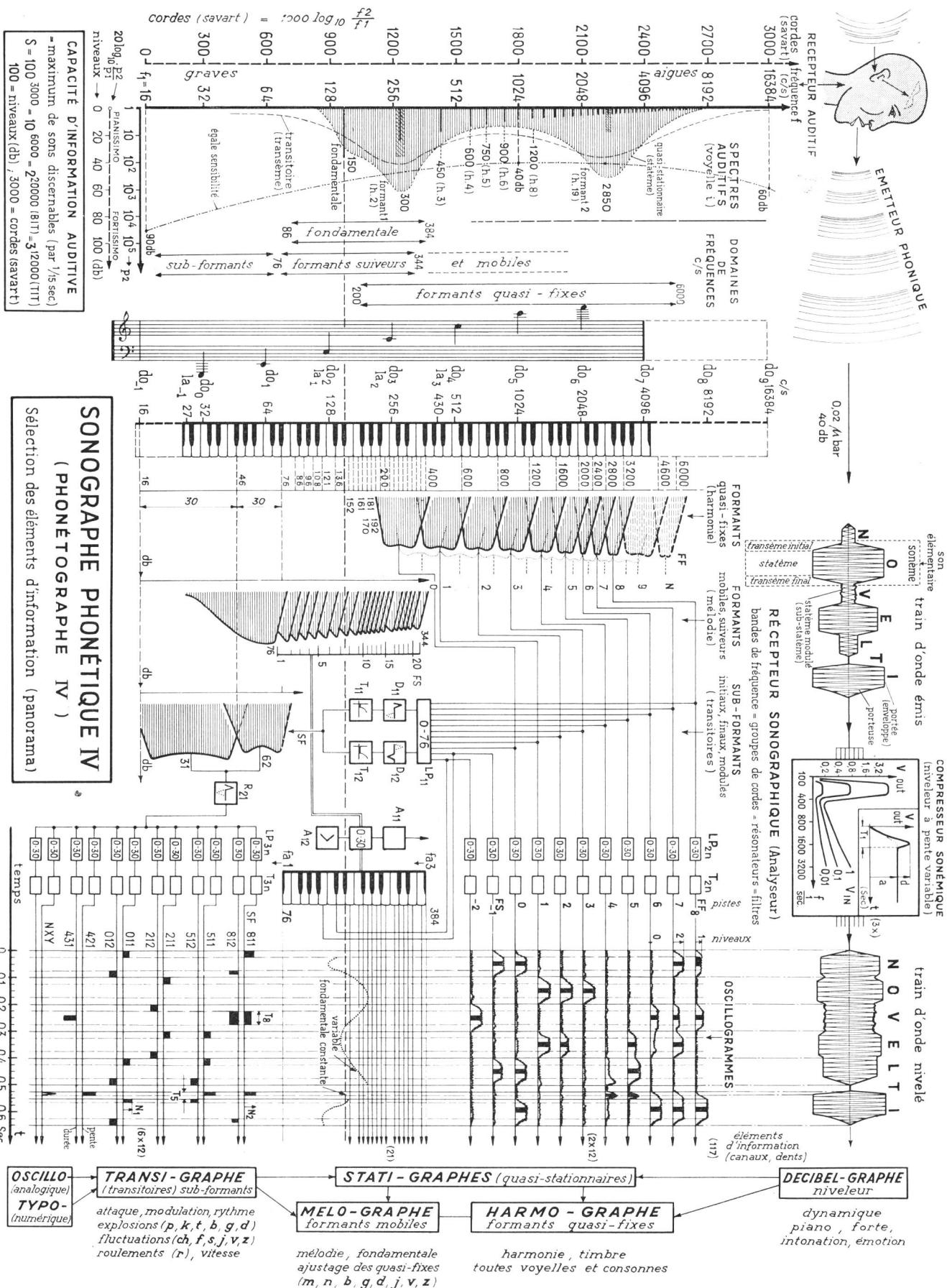


Fig. 1. Sonographe phonétique IV, sélection des éléments d'information

cordes pour obtenir quand même le même nombre S de sons discernables ?

La théorie de l'information donne la réponse suivante :

$$S = 100^{3000} \text{ (savart)} = 10^{6000} \text{ (DIT)} = 3^{12000} \text{ (TIT)} = 2^{20000} \text{ (BIT)}.$$

On peut donc choisir à volonté les spectres d'information équivalents suivants :

soit 3000 cordes à 100 niveaux, nommées savarts, dans le système auditif naturel, à base 100 ;
soit 6000 cordes à 10 niveaux, nommées DIT ou decimal digit, dans le système décimal (base 10) ;
soit 12 000 cordes à 3 niveaux, nommées TIT ou ternary digit, dans le système ternaire (base 3) ;
soit 20 000 cordes à 2 niveaux, nommées BIT ou binary digit, dans le système binaire (base 2).

Dans ce dernier cas, chacune des 20 000 cordes ne pourrait que vibrer ou ne pas vibrer : elle répondrait par tout ou rien, par oui ou non, et elle constitue l'unité d'information ou BIT des techniciens.

Mais c'est le DIT qui est le mieux adapté à nos habitudes de système décimal. On peut d'ailleurs se passer de tout terme savant, en parlant simplement de zéros. Ainsi, quand il s'agit de caractériser un ordre de grandeur élevé, l'homme moderne remplacera avantageusement les termes «milliers», «millions», «milliards», par «trois-zéros», «six-zéros», «neuf-zéros», etc., ce qui lui permet d'effectuer des calculs logarithmiques sans même s'en rendre compte. En effet, quand on multiplie un «trois-zéros» par un «six-zéros», il suffit d'ajouter le six au trois pour obtenir le neuf du «neuf-zéros», ou milliard, résultant. De nos jours, le voyageur devient d'ailleurs logarithmique sans le savoir : il met autant de temps pour aller de son domicile à l'aérodrome, distant de 5 kilomètres, que pour franchir les 500 kilomètres qui séparent celui-ci de la prochaine capitale.

Nous nous excusons de cette digression abstraite. Nous donnerons plus loin un modèle concret de la théorie de l'information : la serrure.

Mais, pour revenir maintenant à la capacité auditive d'information ou nombre de sons discernables par quinzième de seconde, nous savons donc que c'est un six-mille-zéros. Ainsi exprimé, ce nombre est moins difficile à digérer et il finit par devenir très raisonnable quand on songe qu'il contient en puissance toutes les musiques, tous les bruits et tous les langages parlés de notre univers.

Il reste, cependant, immense quand on le compare au nombre total de toutes les particules de l'univers qui est supposé inférieur à un cent-zéros ou au nombre irréalisable des 2^{64} grains de blé demandé par le joueur d'échecs, mais qui ne représente qu'un dix-neuf-zéros.

Après avoir plané sur la capacité générale du récepteur auditif, nous allons essayer de pénétrer dans la jungle des sons pour identifier ses représentants les plus importants, les éléments phonétiques.

2. Les éléments phonétiques isolés

Si chaque élément phonétique, voyelle ou consonne, pouvait être émis individuellement entre deux silences, le problème serait relativement simple. En effet, on connaît déjà depuis Helmholtz le mécanisme de l'émission des voyelles en régime permanent. Résumons, par exemple, celui du «i». Les cordes vocales émettent des vibrations en dents de scie, riches en harmoniques. Si la fondamentale ou fréquence la plus basse est à 150 c/s, soit environ celle du ré₂ dièze, son spectre indique des barres d'amplitudes décroissantes à 300 c/s (ré₃ dièze), 450 c/s (la₃ dièze), 600 c/s (ré₄ dièze) et ainsi de suite. Ces harmoniques excitent certaines résonances buccales, ou cavités «formées» par la bouche afin d'articuler un son déterminé. D'une manière générale, on nomme «formants» les régions de fréquence caractéristiques d'un son. Les formants du «i» sont au nombre de deux, autour de 300 c/s (ré₃ dièze) excité par l'harmonique n° 2, et autour de 2850 c/s (fa₆ dièze), excité par l'harmonique n° 19 et ses voisins. En modifiant la hauteur de la fondamentale masculine ou féminine qui peut varier entre 86 c/s (fa₁) et 384 c/s (sol₃), environ, on déplace beaucoup les numéros d'harmoniques, mais très peu les cavités buccales, de sorte que ce sont environ les mêmes formants, appelés quasi-fixes, qui sont excités par d'autres harmoniques. Leurs régions sont très floues, en raison de la mollesse des parois buccales et des fluctuations inévitables de la hauteur de voix.

De plus, le récepteur auditif distord profondément le spectre émis, d'une part parce que sa courbe d'égale sensibilité favorise certaines fréquences à mesure qu'elles s'élèvent, jusque vers 3000 c/s (sol₆), d'autre part à cause de sa réponse logarithmique en amplitude.

Ainsi, on obtient pour le spectre auditif de la voyelle «i» tenue, la courbe hachurée dans le diagramme de la figure 1. Ses éléments d'information semblent être au nombre de deux seulement (au lieu des 3000 savarts possibles). Ce sont ses deux formants, situés vers 300 et vers 2850 c/s, dont il suffit de diviser l'amplitude maximum en 2 ou 3 niveaux (au lieu des 100 dB possibles). De même, chaque autre voyelle isolée peut être identifiée par un spectrographe assez simple, disposant par exemple de six filtres de formants entre 100 et 3500 c/s.

Le phonétographe, prototype I, reposait sur ce principe, en 1952. En associant à chaque filtre des relais distinguant 2 ou 3 niveaux d'amplitude, on obtenait la discrimination théorique de $2^6 = 64$ ou de $3^6 = 81$ sons différents. Mais, pratiquement, il n'a pu séparer que 12 voyelles (OU, O, A, E, EU, U, E, E, I, ON, AN, IN) et 5 consonnes (L, M, S, CH, F). Il refusait, par exemple, de distinguer les plosives (P, K, T) des fluctuées (S, CH, F), les non voisées des voisées (B, G, D ou Z, J, V), celles-ci des voyelles (U, E, I), les N des M et L, et les O des R. On était encore en plein désordre. Et, surtout, il ne parvenait pas à décomposer un mot en ses éléments phonétiques,

sauf pour quelques combinaisons de voyelles avec trois consonnes (L, S, CH). Pourquoi ?

Il a fallu attendre les prototypes II et III pour découvrir les éléments d'information manquants : les sub-formants en 1956, puis les niveleurs et les suivreurs, maintenant.

3. A la recherche des éléments d'information manquants

a) Les sub-formants

Nous allons essayer de trouver, ensemble, les éléments d'information cachés dans des enchaînements de sons, tels que ceux du langage parlé. Trouver, c'est avant tout éliminer. S'enrichir en réalité, c'est s'appauvrir en possibilités. L'information est un tel processus d'appauvrissement dispensateur de richesses.

Regardons, à titre d'échantillon, le train d'ondes qui véhicule le mot *Novelti*, tel qu'il apparaît sur l'écran d'un oscilloscope, mais stylisé.

Nous y voyons l'onde porteuse, symbolisée par une sinusoïde modulée. Nous en connaissons un peu les composantes quasi stationnaires, dans la partie limitée par des horizontales. En effet, le spectre du «i» nous avait révélé qu'il s'agissait d'un mélange de fréquences comprises surtout entre 100 et 6000 c/s dont les dominantes, appelées «formants», sont très simples.

Par contre, nous ne voyons pas du tout comment on pourrait distinguer le son «t» du son «i» auquel il semble complètement soudé. Cependant, les obliques reliant les horizontales manifestent de petites différences de pente. Alors, pourquoi ne pas essayer d'analyser l'enveloppe du train d'ondes, son onde portée ? «Peine perdue», ont tendance à répondre des spécialistes, «son spectre n'ajoutera que des fréquences très basses, avec des amplitudes très faibles, auxquelles l'oreille est pratiquement insensible».

Pourtant, l'analyseur du phonétographe II a montré en 1956 que les spectres de l'enveloppe, ou plus exactement des variations de formants, pouvaient fournir une nouvelle catégorie de fréquences, située entre 16 et 76 c/s, les «sub-formants», permettant d'identifier les consonnes explosées P, K, T, fluctuées CH, F, S, roulées R, et les transitoires en général.

Cependant, nous n'étions pas au bout de nos peines. Les niveaux des formants étaient faussés par la faiblesse des consonnes relativement aux voyelles, et autres irrégularités. En effet, bien que le V soit aussi important que le E, en tant qu'information, son niveau général est bien plus faible. Comment «démocratiser» ce cortège irrégulier de sons afin de conférer aux formants leurs vraies valeurs ? Il a fallu inventer le compresseur sonémique ou niveleur que nous allons décrire sommairement.

b) Compresseur sonémique ou niveleur

Nous appelons «sonème» tout élément de son à identifier. Le compresseur sonémique est un amplificateur dont le gain est variable en fonction de l'ampli-

tude et de la fréquence, selon deux lois distinctes, de manière à égaliser les niveaux d'un groupe de sonèmes à discriminer.

Par exemple, le compresseur sonémique de voyelles et semi-voyelles obéit aux courbes indiquées.

A 800 c/s par exemple, le niveau d'entrée V_{in} varie de 0,01 à 1, donc au centuple, tandis que le niveau de sortie V_{out} ne varie que de 0,1 à 0,2, donc au double. Ainsi, son taux de compression est $100:2 = 50$. Par contre, à 200 c/s, le niveau de sortie varie de 0,2 à 4, et le taux de compression est tombé à $100:20 = 5$.

On constate que le domaine de la fondamentale, de 100 à 400 c/s environ, ne contribue que peu à la compression, qui est commandée surtout par les harmoniques. Il s'agit donc d'une compression sélective en fréquence. Il faut encore remarquer que les compressions ne se produisent pas instantanément mais qu'elles ont besoin d'un certain temps d'établissement T_1 . D'autre part, il peut se produire un dépassement d', ou – au contraire – une insuffisance momentanée. La technique permettra d'obtenir toute allure de compressions désirées, dans les détails desquelles nous n'avons pas à entrer ici.

Il suffit de constater que notre train d'ondes, qui était très irrégulier à l'émission, semble avoir passé par un laminoir qui confère à chaque sonème le même niveau global, environ.

Etant ainsi débarrassé de ses variations de niveau parasites, le train d'ondes est analysé par une dizaine de filtres de formants, tels que 0 à 8 ou à N, entre 200 et 6000 c/s, qui fournissent, après redressement et filtrage de 0 à 30 c/s, des variations lentes d'amplitude, telles que FF_0 à FF_8 , nommées «formants quasi fixes». On en voit les oscillogrammes schématisés. Nous disposons ainsi d'une première dizaine d'éléments d'information. Il suffit d'en choisir 2 ou 3 niveaux 0, 1, 2 pour identifier les parties quasi stationnaires, nommées «statèmes», des voyelles telles que O, E, I ou consonnes soutenues – telles que N, V, L.

Quant aux sub-formants, on les obtient comme suit : les ondes partielles fournies par les filtres de formants sont redressées et filtrées de 0 à 76 c/s, puis leurs variations initiales D_{11} et finales D_{12} sont elles-mêmes analysées par des filtres de sub-formants, situés autour de 30 et de 60 c/s. Ceux-ci fournissent deux nouvelles dizaines d'éléments d'information, tels SF_{811} ou SF_{812} , dont il suffit encore de choisir 2 ou 3 niveaux pour identifier les parties transitoires des sonèmes, nommées «transèmes» : par exemple, l'attaque des voyelles O, E, I, ou les consonnes plosives P, K, T, ou les fluctuations du V.

La première partie du phonétographe III a été construite sur ces principes, puis expérimentée, entre 1957 et 1959.

Mais les difficultés suivantes se sont alors manifestées :

- Le compresseur sonémique, ou niveleur, modifie les sub-formants. Il aurait fallu prévoir plusieurs

niveleurs, adaptés à divers groupes de voyelles ou consonnes, au lieu d'un seul.

- b) Plus le nivelleur était précis, et plus les écarts des formants, qui ne sont que quasi-fixes, se manifestaient selon la voix du parleur. Il aurait fallu prévoir un correcteur automatique des formants en fonction de la hauteur de voix.
- c) La frappe des touches de la machine à écrire n'était pas synchronisée avec la succession des sonèmes. En effet, leurs phases étaient commandées et auto-répétées par les formants seuls.

Il aurait fallu prévoir une deuxième horloge ou échelle des temps t (sec) mesurant les intervalles entre sub-formants pour obtenir le rythme de succession des sonèmes.

Il était trop tard pour ajouter ces divers organes complémentaires au prototype III, mais il n'y avait pas de problème technique à les prévoir dans le futur phonétographe IV, ici décrit.

Cependant, une autre difficulté, plus troublante, se manifesta :

Impossible de distinguer les consonnes voisées des consonnes sourdes ou bruitées, ni de certaines voyelles – par exemple, le B du P, le V du F, ou le Z du I.

Les parleurs avaient perdu leur voix. Il devait donc exister encore une catégorie d'informations qui avait échappé à nos explorations.

c) *Les formants-suiveurs*

Dans tout cortège de sons, qu'il soit musical ou linguistique, le récepteur auditif a la faculté de distinguer 5 catégories d'éléments d'information : l'intonation, l'harmonie, le bruit, le rythme et la mélodie.

L'*intonation* se trouvait dans la commande automatique du nivelleur, en même temps que les autres variations parasites de niveau. Mais elle servait, pour l'instant, à s'éliminer elle-même de l'information purement phonétique.

L'*harmonie*, par exemple le timbre des voyelles, nous l'avions bien sélectionnée à l'aide d'une dizaine de formants, du moins pour une voix donnée.

Le *bruit*, qui caractérise l'attaque des sonèmes et leurs fluctuations, nous l'avions, également, identifié à l'aide de deux dizaines de sub-formants.

Le *rythme* ? Rien ne nous empêchera de l'ajouter à l'aide d'une horloge de frappe synchronisée par ces mêmes sub-formants.

Mais la *mélodie* ? Elle avait disparu avec la voix.

Après bien des essais infructueux, à l'aide des filtres les plus divers, le phonétographe III a fini par nous donner la clef du mystère : la voix perdue des consonnes explosées ou fluctuées, B, G, D, J, Z, V, se trouve dans une région mobile de fréquences très basses qui suit la fondamentale à une distance de quelques demi-tons. Son amplitude est très faible, mais cependant très énergique comparativement aux autres sonèmes.

Nous avons nommé cette région : formants-suiveurs. Du même coup, elle permettait de réperer la mélodie,

qui brille surtout par son absence, car elle est généralement noyée dans les premiers formants quasi-fixes.

Cette découverte était due à la machine. Un peu comme l'algèbre de *Boole* ou la logistique, qui continuerait à dormir dans quelques bibliothèques si le calculateur automatique, moderne prince charmant, ne l'avait réveillée.

Le phonétographe IV sera donc pourvu d'une vingtaine de filtres dits de «formants mobiles», dont les frontières seront échelonnées entre 76 et 344 c/s. Parmi eux se trouveront automatiquement deux «formants-suiveurs», juste au-dessous de la fondamentale, dont ils exprimeront la mélodie, tout en ajustant les filtres de voyelles, et en fournissant l'information manquante des consonnes.

L'actuel phonétographe III ne possède que deux formants-suiveurs, dont les frontières sont à 110 et 125 c/s. Le parleur doit maintenir sa voix à 150 c/s (ré₂ dièze).

En regardant attentivement le spectre quasi stationnaire du «i», on se rend compte que la fondamentale à 150 c/s ne se distingue pas par sa force, mais par la faiblesse des fréquences immédiatement inférieures. Comparativement, ces régions ont des amplitudes plus grandes au moment de la transitoire, c'est-à-dire de l'attaque du «i». Il est par conséquent plausible que des consonnes explosées ou fluctuées y produisent des phénomènes intéressants.

Une remarque : en cas d'analyse musicale, les formants mobiles peuvent monter parallèlement aux quasi-fixes jusqu'à la limite des 16 000 c/s. En effet, les résonances de divers instruments de musique, tels que la flûte, se déplacent avec la fondamentale.

Le hautbois, par exemple, est reconnaissable aux résonances de son quatrième et de son cinquième harmonique.

Les éléments d'information du langage parlé sont bien plus complexes que ceux d'un seul instrument de musique. Leurs résonances tiennent du violon par les formants quasi fixes, de la flûte par les suiveurs, du tambour par les sub-formants et d'autres encore. C'est tout un orchestre en miniature que nous promenons dans notre bouche.

Cependant, on constate que le sonographe IV aura éliminé la majeure partie des 3000 intervalles de fréquence, des 100 niveaux et des six-mille-zéros de sons possibles. Ses éléments d'information assimilables aux cordes d'un piano ou aux dents d'un peigne, ne sont plus qu'une centaine et ses niveaux deux ou trois. Ils suffisent néanmoins pour alimenter toute une tribu de «graphes» analogiques ou numériques, c'est-à-dire d'oscillographes et d'imprimantes distribuant des renseignements précieux.

Les imprimantes peuvent taper des lettres, des chiffres, des mots tels que des notes musicales ou des phrases entières.

Par exemple, les éléments de commande des niveleurs fournissent les *décibelgraphes*, indiquant la dynamique ou niveau général, entre pianissimo et fortissimo, ou l'intonation ou l'émotion du parleur.

Les *statigraphes*, émanant des parties quasi stationnaires de sonèmes, peuvent se diviser en *harmographes* ou en *mélographes*, selon qu'ils sont rattachés aux *formants mobiles* ou *quasi fixes*. Ils peuvent renseigner, à volonté, sur l'harmonie ou sur la mélodie, sur les voyelles et sur les consonnes soutenues, sur le timbre ou sur la personnalité du parleur.

Les *transigraphes*, émanant des *sub-formants*, caractérisent les consonnes explosives, fluctuées ou roulées, mais aussi le rythme de succession des sonèmes. Ils fournissent donc des *rythmographes*.

La figure 2 reproduit les oscillogrammes du mot *Novelti*, tels qu'ils sont produits par le phonétographe III associé à un oscillographe OED. Nous voyons le train d'ondes d'abord, sans niveleur, puis après avoir traversé celui-ci. On constate que les consonnes sont devenues les égales des voyelles, environ.

Ensuite, nous voyons le résultat de l'analyse par les formants 8 à 0, puis par les suiveurs -1 et -2 qui caractérisent le N et le V, et finalement par quelques sub-formants identifiant le T et complétant le V.

Le phonétographe dispose actuellement de 10 formants quasi fixes, de 2 suiveurs et de 6 sub-formants, soit de 18 éléments d'information ternaires, c'est-à-dire à 3 niveaux, au total – ce qui permet de discerner

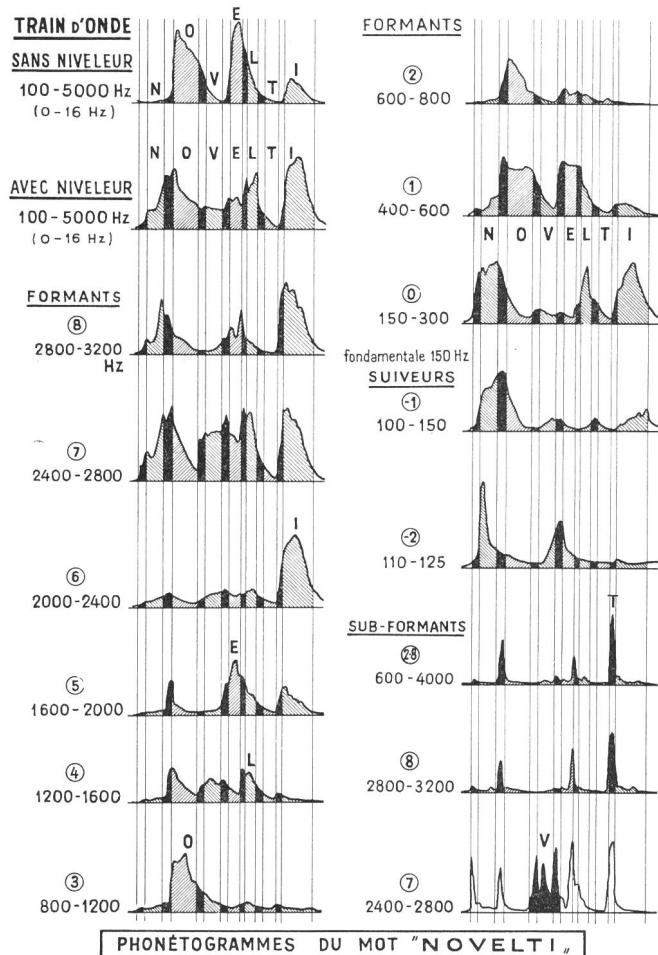


Fig. 2. Phonétogrammes (oscillogrammes) du mot *Novelti*

3^{18} = quatre cent millions de sons différents. Ce chiffre est insuffisant à alimenter les correcteurs. Néanmoins, il semble énorme en regard de la trentaine de touches d'une machine à écrire phonétique. Comment découvrir les quelques combinaisons utiles ?

On est un peu dans la situation d'un propriétaire de 30 clefs qui se trouverait en face de quatre cent millions de serrures sans savoir lesquelles sont les bonnes.

La figure 3 résume les spectres phonétiques et du même coup quelques dizaines ou centaines de combinaisons utilisables. La première colonne concerne les sonèmes liés dans des syllabes, la deuxième les sonèmes isolés. En première approximation, on constate bien que les formants du «i» se situent vers 300 et 3000 c/s, ceux du «a» vers 1000 c/s, que ceux des P, K, T n'existent pas, mais que l'information de ces consonnes réside dans les sub-formants.

4. Sons et couleurs

Nous pouvons faire la tentative d'expliquer d'une façon plus visuelle et concrète le fonctionnement du phonétographe en transposant le monde des sons dans celui des couleurs, car bien des personnes ne savent pas distinguer un ré d'un sol, tandis que chacun connaît la différence entre le rouge et le jaune.

Il existe, d'ailleurs, une correspondance entre le monde des sons et celui des couleurs de même qu'entre les domaines de la musique et de la peinture.

Tous les deux ont l'air d'évoluer du concret vers l'abstrait, c'est-à-dire de s'adresser de moins en moins aux sens et à l'intuition et davantage à l'intellect comme à la raison, par l'intermédiaire de codages de plus en plus subtils. Cependant, ce n'est qu'une illusion chronique, ces termes étant tout relatifs. L'abstrait d'aujourd'hui deviendra le concret de demain. Les classiques actuels sont les mêmes personnages que les révolutionnaires d'hier.

Le même phénomène semble, également, se produire dans le domaine de la science: les langages mathématiques deviennent de plus en plus sténo-graphiques, généraux et condensés tout à la fois. Mais, les générations futures les apprendront dès leur enfance; elles se promèneront avec tout autant d'aisance dans les espaces à n dimensions que les épiciers d'hier parmi n étalages de denrées alimentaires différentes.

Celui qui a compris pour lui-même certaines lois en langage abstrait peut généralement les traduire en images concrètes pour les autres. Sinon, il y a de fortes présomptions qu'il n'ait pas compris.

Examions de plus près l'analogie entre les sons et les couleurs. Les chiffres que nous citerons ne seront que des ordres de grandeur, pour ne pas charger l'exposé. Si l'œil regarde une feuille de papier correspondant au champ visuel, son pouvoir séparateur (linéaire) le long d'une colonne verticale – par exemple – est tel qu'il peut distinguer jusqu'à 3000 raies parallèles environ.

Nº	c/s	ou	ô	o	â	â	a	e	ö	u	ê	é	i	on	en	un	in
	FORMANTS	u	o	ɔ	a	ɑ	a	ə	ø	ʏ	ɛ	e	i	ɔ	ã	œ	ɛ
quasi-fixes																	
9	4600 - 6000	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
8	2800 - 3200	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
7	2400 - 2800	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
6	2000 - 2400	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
5	1600 - 2000	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
4	1200 - 1600	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
3	800 - 1200	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
2	600 - 800	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
1	400 - 600	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
0	150 - 300	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
suiveurs																	
-1	100 - 150	●	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○
-2	110 - 125	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
SUB-FORMANTS																	
2-8	600 - 4000	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
8	2800 - 3200	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
6	2000 - 2400	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
4	1200 - 1600	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
3	800 - 1200	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
1	400 - 600	○	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●

Nº	c/s	l	m	n	r	z	s	j	ch	v	f	b	p	g	k	d	t
	FORMANTS	l	m	n	r	z	s	j	ch	v	f	b	p	g	k	d	t
quasi-fixes																	
9	4600 - 6000	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
8	2800 - 3200	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
7	2400 - 2800	○	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
6	2000 - 2400	○	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
5	1600 - 2000	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
4	1200 - 1600	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
3	800 - 1200	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
2	600 - 800	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
1	400 - 600	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
0	150 - 300	○	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
suiveurs																	
-1	100 - 150	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
-2	110 - 125	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
SUB-FORMANTS																	
2-8	600 - 4000	○	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○
8	2800 - 3200	○	○	○	●	●	●	●	●	●	●	●	●	●	●	●	●
6	2000 - 2400	○	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
4	1200 - 1600	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
3	800 - 1200	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
1	400 - 600	○	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○

SPECTROGRAMMES		Voyelles + consonnes (parleur JDG, fondamentale 150 c/s Formants et sub-formants initiaux, filtres 0-16 c/s)															
sonèmes		○ niveau 0	○ niveau 0	○ niveau 0	○ niveau 0	○ niveau 0	○ niveau 0	○ niveau 0	○ niveau 0	○ niveau 0	○ niveau 0	○ niveau 0	○ niveau 0	○ niveau 0	○ niveau 0	○ niveau 0	○ niveau 0
I liés 4,5/sec.	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
II isolés 1 sec.	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

Fig. 3. Spectrogrammes du phonétographe III

Or, l'oreille peut distinguer 3000 hauteurs de son environ, entre la note la plus grave et la plus aiguë.

Le nombre de couleurs ou teintes que l'œil peut distinguer aisément est d'une centaine entre le rouge et le violet.

L'oreille, elle, peut discerner une centaine de niveaux d'intensités sonores entre le pianissimo d'un murmure et le fortissimo d'une trompette.

La capacité d'information (linéaire) de l'œil, le long d'une colonne verticale, est donc similaire à celle de l'oreille, soit: 100 multiplié 3000 fois par lui-même, ce qui équivaut à 10 élevé à la puissance 6000,

par élément de temps discernable, c'est-à-dire par quinzième de seconde en moyenne pour l'oreille.

Ce chiffre énorme représente le nombre maximum d'images linéaires et de sons entre lesquels on peut établir une correspondance.

Un impressionniste ou un pointilliste qui juxtapose des couleurs franches exploite bien mieux la capacité d'information de l'œil qu'un peintre ancien qui les mélange d'avance. C'est pourquoi un Rembrandt n'atteindra jamais la luminosité d'un Renoir.

Pour expliquer optiquement l'analyse acoustique par le phonétographe, nous établissons un tableau

de correspondance entre la gamme des fréquences acoustiques et celle des couleurs. La fondamentale ou mélodie de la voix humaine pouvant varier entre 80 et 400 c/s est représentée par une vingtaine de teintes rouges allant du foncé au clair.

Ensuite, la gamme des harmoniques, allant de 400 à 6000 c/s, est décomposée en neuf bandes de fréquences dénommées «formants quasi-fixes» qui sont représentés par neuf couleurs de l'arc-en-ciel allant de l'orangé au bleu outre-mer – selon le tableau suivant:

fondamentale mélodie	formants quasi fixes harmonie					
rouge foncé à rouge clair (20 teintes) 80 à 400 c/s	orange 400-600	jaune orange 600-800	jaune 800-1200	jaune citron 1200-1600	vert jaune 1600-2000	c/s
	vert 2000-2400	vert bleu 2400-2800	bleu 2800-3200	bleu outre-mer 4000-6000	c/s	

Si le train d'ondes acoustiques correspondant, par exemple, au mot NOVELTI pouvait se colorer et défiler devant une lunette, il aurait l'allure d'une suite de petits nuages éclairés sur un fond de ciel nocturne.

En effet, on peut transposer optiquement – comme suit – l'émission phonétique: les cordes vocales émettent une lumière blanchâtre, contenant donc un mélange de toutes les couleurs. Or, les cavités buccales, et autres, «formées» par le parleur pour articuler un mot, colorent cette lumière blanche par suite de résonance. Cette coloration correspond aux formants quasi-fixes déterminant le timbre particulier de chaque voyelle ou consonne quasi stationnaire.

Le O serait orangé, le V bleuté, par exemple.

L'oreille est capable de disséquer naturellement chaque son complexe et d'y reconnaître les composantes de sons purs, contrairement à l'œil qui a besoin d'un auxiliaire artificiel, tel qu'un prisme de spectroscope, pour reconnaître les couleurs pures dans un mélange. Ainsi, pour «voir» la physiologie du phonétographe, qui imite l'oreille humaine, nous le considérons comme un prisme décomposant les couleurs impures en leurs éléments constitutifs.

Le violacé du N serait formé d'autres raies spectrales que le violacé du I.

Les spectres quasi stationnaires des éléments phonétiques, dont le nombre reste inférieur à une centaine, présentent une certaine analogie avec quelques spectres visibles parmi ceux des éléments atomiques.

D'autre part, un train cohérent d'onde lumineuse, émanant d'un atome dure, par exemple, un centième de microseconde et se propage à la vitesse de 300 000 km/sec. Sa longueur observable est donc de trois

mètres. Or, une période d'onde acoustique, émanant d'un parleur dont la hauteur de voix se trouve à 110 c/s, dure 9 millisecondes et se propage à la vitesse de 330 m/sec. Sa longueur observable est donc de trois mètres, également.

5. Clefs phonétiques et serrures frontales

Reste à savoir par quels moyens les éléments d'information peuvent déclencher les touches désirées de la machine à écrire ? Ces moyens font intervenir des techniques de codeurs et de matrices qu'il serait fastidieux de décrire ici en détail. Toutefois, on peut imaginer un modèle qui concrétise non seulement ces techniques, mais encore l'analyse combinatoire, la théorie de l'information et diverses notions abstraites de la physique, telles que l'entropie, le calcul vectoriel et la transformation de coordonnées: c'est la serrure frontale.

A droite de la *figure 4*, nous voyons – en coupe – 3 serrures V, E, L placées l'une derrière l'autre dans le même bloc, sans se gêner. Elles sont constituées par 3 tiges carrées verticales qui sont transpercées par 4 axes ronds horizontaux. Chacun de ces axes transversaux est segmenté en 7 tronçons dont les noirs seuls bloquent les tiges verticales, parce qu'ils y sont partiellement engagés, tandis que les blancs sont neutres – vu qu'ils sont entièrement noyés dans les tiges.

Regardons, par exemple, le fonctionnement de la première serrure V: sa tige carrée verticale est bloquée par deux tronçons marqués en noir. Le tronçon supérieur est un peu moins engagé dans la tige que l'inférieur. Maintenant, examinons la clef V: elle comprend 4 dents de longueurs inégales. Les deux plus courtes, blanches, sont au niveau neutre 0.

Imaginons, maintenant, que la clef se déplace vers la droite: pour commencer, elle repoussera la saillie de la barre générale qui comprimera 3 ressorts. Ceux-ci cherchent à faire monter les 3 tiges verticales qui, pour l'instant, sont encore toutes bloquées par les segments noirs. En continuant à se déplacer vers la droite, jusqu'à buter contre la cloison, la clef V enfonce les 2 axes qui font face à ses 2 dents noires, et les 2 segments noirs intérieurs vont se trouver exactement noyés dans la tige verticale V, qui, ainsi débloquée, jaillira vers le haut et imprimera la lettre V sur une bande de papier.

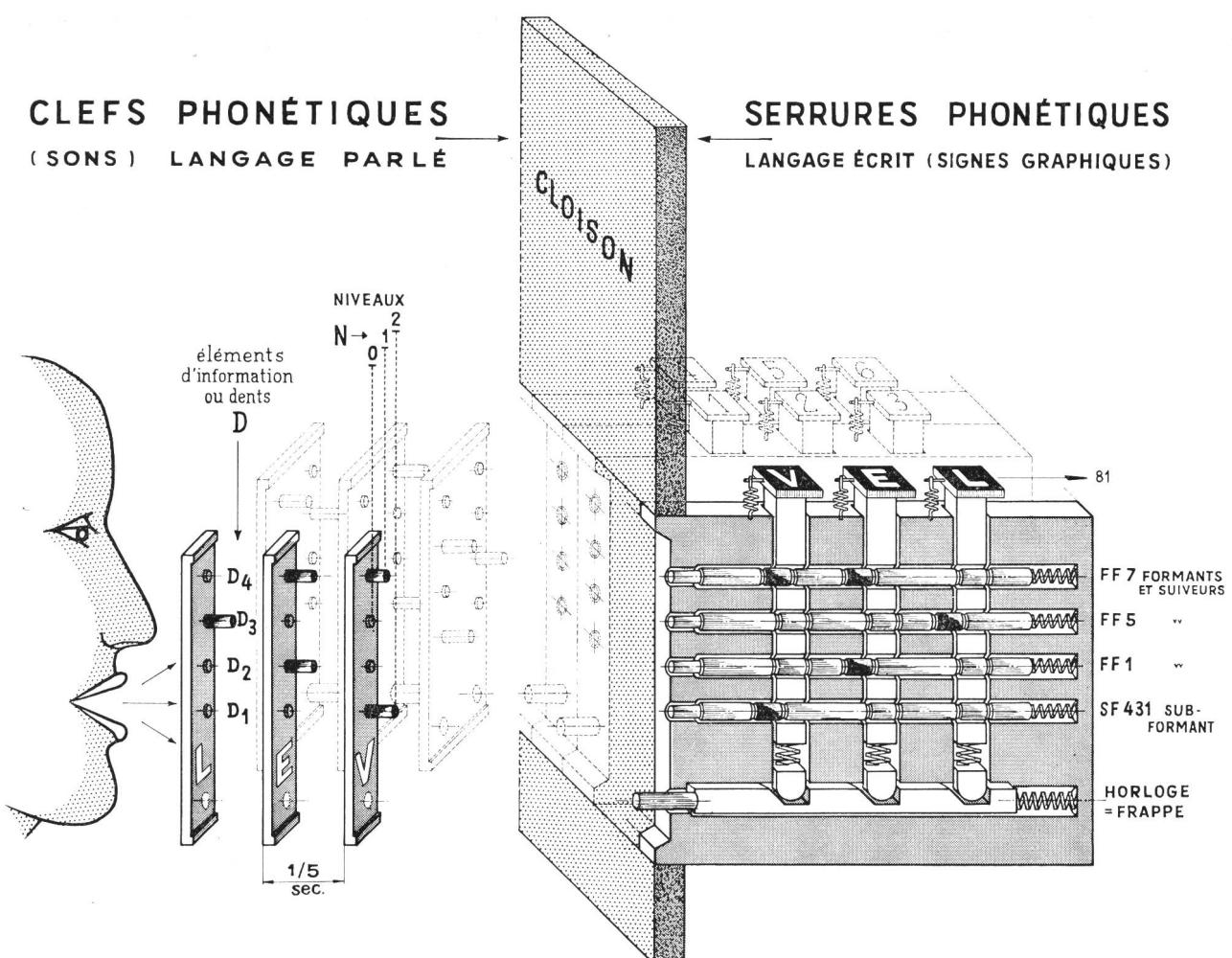
Si la clef se retire, tout revient à sa position initiale de repos sous l'effet des divers ressorts, et tout est prêt à recommencer ou à recevoir la visite d'autres clefs, telles que E ou L, dont chacune n'actionnera que sa tige verticale correspondante E ou L.

Il est facile de calculer le nombre maximum de serrures ou de clefs différentes.

Chaque dent peut avoir l'une des 3 longueurs 0, 1, 2 et se combiner avec chacun des 3 niveaux de chacune des autres dents. Il faut donc multiplier le chiffre 3 autant de fois par lui-même qu'il existe de dents, c'est-à-dire 4 fois, pour obtenir le nombre maximum $3^4 = 81$ serrures qui peuvent être placées l'une

CLEFS PHONÉTIQUES
(SONS) LANGAGE PARLÉ

SERRURES PHONÉTIQUES
LANGAGE ÉCRIT (SIGNES GRAPHIQUES)



NIVEAUX	N	2	2	2
ÉLÉMENTS, DENTS	D	4	$+$	6
SONS, SIGNES	$S = N^D$	$2^4 = 16$	\times	$2^6 = 64$

$$= 2^{10} = 1024$$

NIVEAUX	N	3	3	3
ÉLÉMENTS, DENTS	D	4	$+$	6
SONS, SIGNES	$S = N^D$	$3^4 = 81$	\times	$3^6 = 729$

$$= 3^{10} = 59049$$

EXEMPLES DE CAPACITÉS D'INFORMATION		OREILLE HUMAINE	ALPHABET TÉLÉSCRIPT.	PHONÉTOGRAPHIE III	PHONÉTOGRAPHIE IV	SONOGRAPHIE IV	SERRURES USUELLES	SERRURE FRONTALE	OEIL HUMAIN *	CERVELET (REFLEXES)	CERVEAU (INTERCONN.)
niveaux	N	100 (db)	2	3	3 (2)	3 (2)	5	3	100 (db)	2	2
éléments ou dents	D	3000 (savart)	5	18	48	96	5	12	10^6	100.000	10^{10}
sons, signes, clefs, serrures (max discernables)	S	N^D 100 3000	2^5	3^{18}	3^{48}	3^{96}	5^5	3^{12}	$100^{(10^6)}$	$2 \cdot 100.000$	$N^{(D^2 \cdot 2^{10^{20}})} = 10^{(3 \cdot 10^{19})}$
unités d'information	B	$10^{10} \times 10^{20}$	32	$4 \cdot 10^8$	$8 \cdot 10^{22}$	$6 \cdot 10^{45}$	3125	531441	$10^{2 \cdot 10^6}$	$10 \cdot 10.000$	
par élément de temps	BIT	20000	5	28	76	152	$\sim 2^{11}$	$\sim 2^{19}$	$2 \cdot (6 \cdot 10^6)$	$2 \cdot 100.000$	
par seconde	sec.	1/15	1/5	1/5	1/15	1/15	—	—	6.10 ⁶	100.000	10^{20} (BIT)
tempo équival. à oreille/sec.	BIT	200	35	4	2	—	—	—	3.10 ⁷	—	—
min.		1/60									

* estimation, image = 1/5 sec.

Fig. 4. Clefs et serrures phonétiques

derrière l'autre, dans notre bloc, sans se gêner et sans se répéter.

Si nous disposons de 2 niveaux seulement au lieu de 3 par dent, le nombre maximum de serrures différentes tomberait à $2^4 = 16$. Il faudrait alors rajouter des dents.

Une dent à 2 niveaux s'appelle unité d'information (binaire) ou BIT. Une dent à 3 niveaux est un TIT, ou élément d'information ternaire.

Faisons, maintenant, un effort d'imagination: le bloc s'appelle: la matrice du phonétographe III. Il

contient 30 serrures verticales nommées éléments phonétiques, qui sont bloqués par les tronçons noirs de 18 axes transversaux nommés formants et sub-formants. D'autre part, chaque clef possède 18 dents à 3 niveaux, et l'orchestre en miniature dissimulé dans notre bouche est capable de cracher à volonté 30 de ces clefs, au rythme de plusieurs par seconde: chaque fois qu'une de ces clefs rencontrera sa combinaison correspondante dans le bloc, elle libérera une tige, et une seule, qui frappera la lettre désirée.

Le modèle imaginé pourrait d'ailleurs se réaliser un

jour. En effet, il permettrait de fabriquer des matrices optico-mécaniques d'un format très réduit, applicables au phonétographe, à des ordinateurs ou autres automates.

Regardons le tableau qui confronte quelques capacités d'information, et que nous pouvons traduire maintenant en termes concrets.

L'oreille humaine contient un six-mille-zéros de serrures dont chacune peut être ouverte par une clef acoustique possédant une combinaison déterminée de 3000 dents et de 100 niveaux.

Le télémimeur se contente de 32 serrures qui peuvent être ouvertes par des clefs minuscules à 5 dents et 2 niveaux.

Quant au phonétographe III, nous avons vu qu'il comprend virtuellement 400 millions de serrures, parmi lesquelles seules une trentaine sont utilisées pour discerner les éléments phonétiques, et que chaque clef correspondante, crachée par l'orchestre buccal, possède jusqu'à 18 dents à 3 niveaux.

On est frappé par le gaspillage d'une telle machine en regard de l'économie d'un téléimprimeur. C'est que le phonétographe est à mi-chemin entre la vie et la machine. Or, la prodigalité de la vie est démesurée. Il suffit de songer aux essaims de pollens qu'elle jette à tout vent pour féconder une seule fleur.

Pour nous reposer de ces chiffres vertigineux, retournons aux serrures de sûreté usuelles qui ferment nos portes: elles ont, généralement, des clefs à 5 dents et à 5 niveaux admettant 3125 combinaisons, au total. Quant à la serrure frontale dont vous pourrez voir quelques prototypes expérimentaux, elle peut – par exemple – posséder des clefs à 12 dents et 3 niveaux permettant environ 500 000 combinaisons.

Pour terminer, nous pouvons avoir la curiosité de connaître, encore, les capacités d'information suivantes :

Celle de l'œil est, au moins, cent fois plus grande que celle de l'oreille, car l'œil dispose d'une dimension de plus : la surface, au lieu de la ligne, ce qui contribue aux avantages de l'écriture.

Quant au cerveau, le nombre de combinaisons qu'il peut contenir dépasse l'entendement de son propriétaire. Tout se passe comme si notre bloc de serrures avait une dimension de plus, en profondeur, et comme si les dents des clefs avaient des niveaux variables, interconnectés. *Pythagore* disait déjà que les différences de qualité reposent sur des différences de quantité. Le nombre de serrures assimilable au cerveau est tellement grand qu'on ne peut même pas l'exprimer directement en milliards-de-zéros. On est obligé de l'écrire par une cascade de puissances, ou de prononcer le terme compliqué de dix-puissance-dix-neuf-zéros.

D'autre part, on peut considérer les langages alphabétiques comme des ensembles de serrures et de clefs, constitués par des racines linguistiques, des mots, des phrases, des pages et des livres en nombres pratiquement illimités. Mais chaque dent de ces clefs n'a qu'une trentaine de niveaux, figurant les

éléments phonétiques facilement mémorisables par quiconque.

La puissance extraordinaire de l'écriture alphabétique provient de ce qu'elle est ainsi capable d'interconnecter, de condenser et d'accumuler les principales pensées de tous les cerveaux présents et passés.

6. Phonétogrammes de mots fournis par le phonéto-graphe III/2

Nous cédon-s la parole – ou, plutôt, la plume – au personnage central de notre conférence, le *phonéotraph*. Il transcrira quelques paroles enregistrées sur magnétophone en utilisant un alphabet phonétique français, puis international.

medaamm éé munnéessiioö vva vvéérréé munininteeenow la pääärolle
deevenkärr hääkkäliiow vvv ässiiissté aa la preemmerriiässii
métamooorrffffooz ootoommäätäitiike du llemangäjj päääärillé
a llemangäjj éekrrii pääärlle fonnnéetogrääff proototiiipe
trronwaaää rrééggllihääjänn nnnuunuméero dö séchamattiiow manglëe
merd gëësts ééekchähn iitiilinc liiz prronouff of paunouding
asalleemaa vvoort viirrd taat

medaamm ee mmmeessiiigø vvu vveeserree mm̄s̄ēt̄ēt̄ēnn̄d
la padarille dævænn̄lir daakss̄i1686 vu dassiisst̄e aa la
preæmm̄fieeser metmæoɔrrrrfføzz œst̄oɔmm̄mædæt̄i ky d 11āād̄gaa33
padarille d 11āād̄gaa33 rriiç̄i paad l1la fonneet̄grøadaf
prøt̄st̄iløe truuuwa rræs̄glllaaḡ nnyyyymmerø dø
eød̄d̄t̄i1686 døglee werd geest̄s eek̄s̄jññ iitiilinng iiz
pruuuuff of puuudig aalaññd̄ vuort viirrd taat̄

Fig. 5. Phonétogrammes transcrit par le phonétographe III, réglage n° 2, en alphabets phonétiques français et international.

Vous avez pu constater que le phonétographe est approximativement en état de métamorphoser la parole en écriture phonétique, mais que les lettres fournies sont souvent répétées et que le parleur doit articuler chaque mot d'une manière lente et monotone. En effet, le prototype III n'est pas encore pourvu d'automatismes d'adaptation au parleur. C'est, au contraire, celui-là qui doit s'adapter à la machine: il doit se synchroniser avec un rythme rigide qui découpe le temps à raison de six tranches par seconde, il doit contrôler la hauteur et l'intensité de sa voix. Le phonétographe III n'existe que depuis l'hiver dernier et le temps mis à notre disposition n'a pas suffi pour optimaliser ses réglages et sa logique. Il faudrait encore quelques mois.

On lui adjoindra une vingtaine de filtres interchangeables pour compléter la gamme des formants suivreurs, qui ne sont qu'au nombre de deux, actuellement.

Alors, le prototype III sera utile à des laboratoires de sonométrie, de phonétiques et autres. Il servira à des recherches statistiques sur un grand nombre de parleurs et facilitera la réalisation du prototype IV.

7. Phonétographe IV

Le phonétographe IV, dont les plans généraux sont élaborés, sera le premier prototype industriel. Il possédera, par exemple, quarante-huit éléments d'information, dont trente-six formants et douze sub-formants à trois niveaux chacun. Il pourra donc discerner $3^{48} = 8.10^{22}$ sons différents par quinzième de seconde. Il sera pourvu de tous les automatismes d'adaptation au parleur, mais la miniaturisation lui permettra de ne pas excéder les dimensions d'un frigorifique usuel.

La machine de bureau simple donnera un texte phonétique sans répétitions, facilement lisible sans long apprentissage, tel que :

v ou v è r é l a p a r o l a d e v e n i r a k s i o n

Mais le phonétographe IV sera muni d'une mémoire permettant d'en extraire, en tout temps, bien d'autres catégories d'informations. En lui adjoignant des oscilloscopes ou des imprimantes spéciales, on obtiendra à volonté des mélographes, des décibelgraphes, des rythmographes, des transigraphes, des syntaxographes, des identigraphes, etc.

A titre d'exemple, le texte imprimé pourra mimer automatiquement la mélodie du parleur, par décalage conforme des lettres :

v v ou ou v v è r r é é l l a a a a p a a r r o o o o l l

En variante, les décalages progressifs de lettres peuvent renseigner sur leurs accentuations.

Mais, on peut aussi prévoir des décalages globaux de mots, en fonction de l'intonation du parleur, et qui donnent des informations grammaticales. Par exemple, dans la phrase :

v ou v è r é l a p a r o l a d e v e n i r a k s i o n

Les mots décalés «véré» et «devenir» se désignant comme verbes.

Mais le phonétographe peut aussi donner des informations personnelles sur le parleur lui-même, sur son caractère, sur sa santé, sur son identité et autres.

8. Automation et «phonaction»

Le phonétographe IV peut servir de base à un nouveau domaine technique qu'on peut appeler «phonaction» et qui complétera l'automation actuellement en plein développement.

Quelque part, il faut toujours une intervention humaine. Jusqu'à présent, c'étaient des doigts qui pesaient sur des boutons ou sur des touches. Bientôt, il suffira de parler non seulement pour actionner une machine à écrire mais aussi pour perforer les cartes d'une calculatrice, pour trier des objets, pour appeler

un ascenseur, pour allumer une lampe, pour ouvrir une porte, pour traduire un mot, voire une phrase, d'une langue dans une autre, etc.

Une fois le problème de la traduction automatique résolu, on pourra du même coup transformer l'écriture phonétique en orthographe, par adjonction d'une calculatrice.

Le phonétographe peut aussi servir à la lecture ou au dépouillement optique de figures, de textes ou de photographies, telles que celles qui sont fournies par des chambres à bulles ou de Wilson.

Il suffit qu'un rayon lumineux balaye l'image à explorer à une cadence telle qu'il en résulte des fréquences analysables selon des procédés analogues à ceux du phonétographe.

Mais le domaine de la «phonaction» englobe encore d'autres applications. Ainsi, on pourra effectuer la téléphonie à longue distance sur bande étroite, réduisant considérablement les frais de la conversation. Un central téléphonique pourra donner des textes lisibles de messages téléphonés.

Ce nouveau domaine d'activité scientifique et industrielle semble convenir particulièrement bien à un pays petit, mais développé, comme la Suisse. Il serait recommandable de créer un «Laboratoire central de phonaction».

On peut s'attendre à ce que la «phonaction» ait des incidences sur diverses sciences morales, qu'elle incitera à devenir de plus en plus exactes et à se rattacher à la physique: en effet, la machine ne comprend que ce qui est rigoureusement logique.

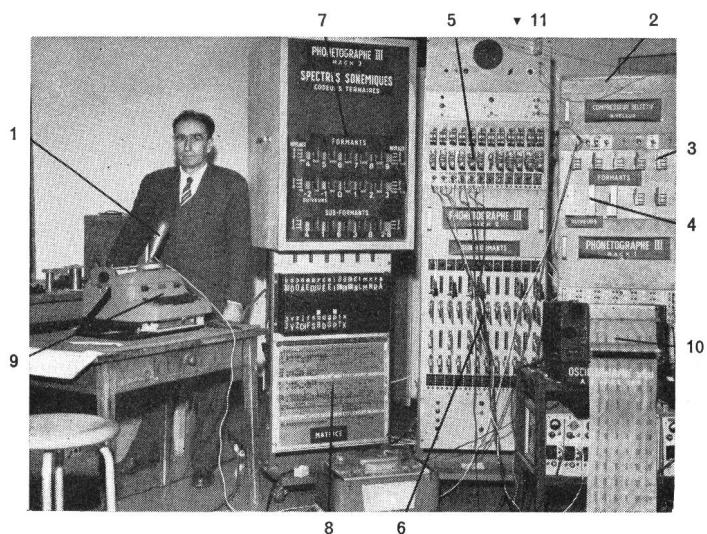


Fig. 6. Phonétographe, prototype III

- 1 Microphone
- 2 Compresseur sélectif ou niveleur sonémique
- 3 Filtres passe-bande de formants quasi fixes
- 4 Filtres passe-bande de formants-suiveurs
- 5 Filtres passe-bas de formants
- 6 Filtres passe-bande et passe-bas de sub-formants
- 7 Codeurs ternaires et spectroscopie sonémique
- 8 Matrice et chaînes logiques
- 9 Machine à écrire électrique input-output
- 10 Oscillographie électro-dynamique OED, à 6 voies SED (graphes analogiques)
- 11 relais électro-dynamique RED, commutant le niveau sonémique 2

Elle oblige l'homme à penser et à agir logiquement. Ainsi, elle pourrait amener progressivement des révisions de l'orthographe, de la phonétique, de la linguistique et même des sciences sociales et du droit.

Le monde humain est dirigé par ces petits nuages évanescents nommés «paroles», mais qui sauront

de mieux en mieux se muer automatiquement en actions logiques.

Rappel de la publication précédente: Bulletin Technique PTT n° 2/1957, p. 41-59.

Adresse de l'auteur: Jean Dreyfus-Graf, ing. dipl., 5, avenue Grenade, Genève.

Verschiedenes - Divers - Notizie varie

Statistique téléphonique mondiale au 1^{er} janvier 1960

31:654.15(100)

Chaque année, l'American Telephone and Telegraph Company (ATT), nous fait parvenir les résultats de la vaste enquête qu'elle entreprend auprès des administrations et des compagnies téléphoniques du monde entier. Il y a quelques semaines, nous avons reçu les statistiques arrêtées au 1^{er} janvier 1960. Bien qu'une année se soit écoulée depuis ce dénombrement, il est intéressant de consacrer quelques instants à ces chiffres: la progression constante de notre pays dans le domaine des télécommunications ne doit pas nous empêcher de faire le point, de suivre aussi le développement accompli à l'échelon mondial.

Répartition des postes téléphoniques par continents

Au 1^{er} janvier 1960, il y avait 133,6 millions de postes téléphoniques dans le monde. Par rapport à fin 1958, l'augmentation est de 7,05% et si nous comparons avec les chiffres de 1950, nous constatons que la progression a été de 90% en 10 ans.

Avec près de 116,5 millions de postes téléphoniques, l'Amérique du Nord et l'Europe possèdent 87,2% des téléphones installés à la surface de la terre. Il faut pourtant souligner que l'Asie a, en 10 ans, plus que triplé le nombre des postes téléphoniques installés sur son territoire et que pour la même période, les téléphones africains ont passé de 0,8 à 1,9 million.

La densité mondiale téléphonique, c'est-à-dire le nombre des postes pour 100 habitants, est de 4,6. C'est en Amérique du Nord que la densité est la plus élevée (38,7); ce continent est suivi par l'Océanie (18,7) et l'Europe (7,0).

Continents	Postes téléphoniques au 1.1.1960	En % du total	Pour 100 habitants
Amérique	80 190 300	60,0	Nord 38,7
			Centre 1,5
			Sud 2,3
Europe	40 340 900	30,2	7,0
Asie	8 110 000	6,1	0,5
Océanie	3 054 300	2,3	18,7
Afrique	1 904 500	1,4	0,8

L'automatisation a continué de progresser dans toutes les parties du monde: l'Asie est toujours au dernier rang avec 66,5% seulement des stations desservies automatiquement.

L'exploitation téléphonique est, à l'échelon mondial, principalement confiée à des compagnies privées (67,1%). Ce sont des organismes gouvernementaux surtout qui assurent le service en Afrique (98,3%), en Océanie (92,7%), en Europe (83,2%) et en Amérique du Sud (51,5%).

Répartition des postes téléphoniques par pays

On compte plus de 1 million de postes dans 17 pays. Les Etats-Unis sont de loin les plus gros consommateurs de téléphones; les 16 autres «millionnaires» ne totalisent que 49,3 millions de postes téléphoniques, alors qu'à eux seuls, les Etats-Unis en ont près de 71 millions!

Tableau I.

Pays	1.1.1960	1.1.1950	Augmentation en % par rapport à 1950
Argentine	1 244 133	723 837	71,9
Australie	2 122 251	1 066 385	99,0
Belgique	1 084 594	649 555	67,0
Canada	5 439 023	2 699 612	101,5
Danemark	1 019 582	680 703	49,8
Espagne	1 641 395	606 066	170,8
Etats-Unis	70 821 000	40 709 000	74,0
France	4 084 843	2 318 673	76,2
Grande-Bretagne (Royaume Uni)	7 848 000 ¹	5 159 436	52,1
Italie	3 517 908	1 118 832	214,4
Japon	4 864 858 ¹	1 681 279	189,4
Pays-Bas	1 500 693	662 000	126,7
Rép. dém. allemande ..	1 237 796	manque	—
Rép. féd. d'Allemagne ..	5 516 226	2 183 000	152,7
Suède	2 637 336	1 591 473	65,7
Suisse	1 562 360	845 471	84,8
URSS	4 022 633	manque	—
¹ 31. 3. 60			

Parmi les «viennent ensuite», il faut citer la Pologne, dont le nombre des postes téléphoniques a augmenté, pour la période de 1950 à 1960, de 244,2%; le dénombrement de 1950 donnait 236 000 postes et celui de 1960, 812 538 téléphones.

En plus de la Pologne, deux pays atteindront certainement le cap du million lors des prochaines statistiques: le Brésil (964 013 postes à fin 1959) et la Tchécoslovaquie (936 099 postes au 1.1.1960).

Depuis 1950, six nouveaux pays ont pris place parmi les «grands» du téléphone. Pour l'Espagne, l'augmentation a été de 170,8%, alors qu'elle a été de 126,7% pour les Pays-Bas, de 84,8% pour la Suisse, de 71,9% pour l'Argentine, de 67,0% pour la Belgique et de 49,8% pour le Danemark.

Au sein des anciens «millionnaires», le développement est aussi remarquable par rapport à 1950, tout particulièrement pour l'Italie (214,4%), pour le Japon (189,4%) et pour la République fédérale d'Allemagne (152,7%).

Les chiffres des pays de l'est européen sont communiqués, semble-t-il, depuis quelques années aux statisticiens américains. Le recensement officiel effectué en Union soviétique en 1936 indiquait 861 181 postes téléphoniques; en 25 ans, le nombre des téléphones a donc presque quintuplé. Par rapport à 1958, cette augmentation est de 13%; cette partie du monde progresse aussi dans le domaine téléphonique, bien que la densité demeure encore extrêmement faible (1,91).

Au point de vue densités nationales, nous obtenons, pour les pays comptant plus de 1 million de postes téléphoniques, le tableau suivant: