

REVUE DE STATISTIQUE APPLIQUÉE

J.-P. BENZECRI

Sur l'analyse des matrices de confusion

Revue de statistique appliquée, tome 18, n° 3 (1970), p. 5-63

http://www.numdam.org/item?id=RSA_1970__18_3_5_0

© Société française de statistique, 1970, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

SUR L'ANALYSE DES MATRICES DE CONFUSION

J.-P. BENZECRI

Professeur de Statistique à la Faculté des Sciences de Paris

D'un objet matériel simple, on peut donner une description mathématique complète, et précise comme une épure ; on peut encore user du langage commun. Ces descriptions ne mettent pas toujours à leur juste place les traits qui, reçus par les sens, ont le plus d'importance pour le sujet qui considère l'objet. A plus forte raison est-il difficile de découvrir par la seule réflexion, sans expérimentation, les caractères par lesquels nous distinguons les objets complexes. Ces caractères définissent, en quelque sorte un espace, où nous situons les choses : la relation de voisinage, la topologie, voire la métrique (les distances), propres à cet espace, ne donneraient-elles pas une expression géométrique précise à la notion confuse de similitude ? Chercher dans le concret et aussi dans l'abstrait, les dimensions de l'esprit, c'est une des premières tâches de la psychologie. Dans cette recherche, l'analyse statistique qui d'un tableau de nombres vise à extraire un ordre spatial, a certes sa place : et nous proposons ici quelques exemples d'analyses de ces tableaux particuliers que sont les matrices de confusion.

Selon une conception simpliste l'acte psychologique élémentaire est décrit par un couple $(s - r)$, où r est la réponse fournie par le sujet au stimulus s . Il est difficile de fonder sur ce schéma une analyse féconde de l'activité naturelle d'un homme ou même d'un animal accomplissant une tâche qui l'intéresse. Mais en contraignant les sujets d'une expérience à répondre à des signaux, le psychologue n'en remplit pas moins des tableaux de paires $(s - r)$, dont l'analyse révèle quelque chose de l'âme... Deux stimuli s et s' qui sont pour le sujet, semblables, recevront de celui-ci des réponses dont les distributions statistiques seront voisines (e.g. si le sujet a appris à répondre r à s , il répondra aussi r à s' voisin de s). Et, de même, deux réponses semblables résulteront d'un même stimulus avec des fréquences conditionnelles peu différentes. Ici l'on parle communément de généralisation de la réponse, là de généralisation du stimulus. (Sans doute vaudrait-il mieux dire généralisation sur ; car ce n'est pas le stimulus lui-même ou la réponse qui reçoit généralisation, mais l'attitude du sujet quant à l'un ou l'autre). Quoiqu'il en soit du protocole expérimental conçu avec plus ou moins d'ingéniosité pour maintenir les sujets en éveil et faire jouer sélectivement telles de leurs facultés, on aboutit en général à un tableau de correspondance k_{SR} : S étant l'ensemble des stimuli utilisés, R celui des réponses permises, $k(s, r)$ est le nombre de fois que, dans une expérience avec un ou plusieurs sujets, le stimulus s a reçu la réponse r . L'analyse du tableau de correspondance k_{SR} fournit pour les ensembles S et R une représentation géomé-

trique (par exemple plane, si l'on extrait deux facteurs), expression concise des similitudes entre les profils des lignes et des colonnes de ce tableau.

Le jeu simultané de deux ensembles distincts S et R donne à l'expérience plus de confusion que de richesse : souvent faute de savoir ce qui provient de la structure de chaque ensemble, on ne peut conclure ; ou encore le lien artificiel mis par l'expérimentateur entre S et R afin de contraindre le sujet à répondre, brouille si bien les structures naturelles de ces ensembles qu'aucune interprétation ne se suggère pour les résultats de l'analyse. C'est pourquoi on s'intéresse surtout aux tableaux carrés $S \times S$ (que ces tableaux satisfassent ou non à la condition de symétrie : $k(s, s') = k(s', s)$). Ces tableaux carrés sont appelés communément : matrices de confusion, comme si $k(s, s')$ était le nombre de fois que s a été pris pour s' : ce qui n'est pourtant pas en général le cas, les paires (s - s') dénombrées dans les expériences étant produites par des jeux complexes dont on verra ci-dessous des exemples. On peut encore étudier les similitudes entre éléments d'un ensemble S par des expériences qui ne conduisent pas au dénombrement, direct ou indirect de paires (s, s') : par exemple, on demandera à des sujets d'exprimer directement par un nombre la similitude entre s et s'. Nous parlerons peu ici (cf. toutefois 1 - 3) de ces expériences que nous entendons traiter avec l'analyse statistique des proximités.

Ailleurs, nous avons considéré l'analyse statistique (distributionnelle) d'un tableau général $S \times R$ (cf. Benzécri 1966). Ici, nous voulons rassembler la quasi-totalité des analyses de correspondances faites, (à notre connaissance), sur des tableaux de couples de stimuli ou de réponses ; en fait, il s'agira toujours (sauf au n° 8) de tableaux carrés, tableaux que nous sommes convenus d'appeler matrices de confusion. Aux exemples effectivement traités, nous avons joint quelques projets : espérant que des chercheurs s'ingénieront à remplir les tableaux que nous voudrions voir analyser...

Les expériences sont présentées en ordre de complexité croissante. Les premiers exemples sont surtout l'occasion d'exposer les difficultés propres aux recherches psychologiques. Comme la structure de l'ensemble des stimuli y est très simple et que les tableaux sont de petite taille, l'analyse factorielle ne peut que faire la preuve de sa propre efficacité, en confirmant des résultats attendus. Mais dans les exemples les plus complexes qu'on ait traités, l'analyse donne déjà des représentations intéressantes qui n'étaient pas totalement prévisibles ; et, ce qui est mieux, elle suggère d'explorer des voies nouvelles de recherche.

Nous nous sommes partout appliqués à faire comprendre le problème psychologique avec concision mais sans ambiguïté. D'une part, en effet, sans ces explications l'examen des graphiques eût vite été fastidieux. D'autre part le statisticien, qui analyse des expériences très diverses, ne peut faire honneur aux méthodes qu'il utilise que s'il est averti des problèmes méthodologiques de toutes les sciences.

1 - UN ENSEMBLE DE HUIT STIMULI COLORES

La matrice de confusion dont nous rapportons ici l'analyse, provient d'une expérience d'apprentissage (cf. Benzécri & Richard, 1964) que nous décrirons d'abord brièvement.

1.1 - On présente à des sujets huit couleurs, projetées grâce à des filtres quasi monochromatiques sur un écran blanc ; les sujets doivent apprendre à associer aux couleurs les boutons d'un clavier, disons les entiers de 1 à 8, pour omettre les détails mécaniques. On procède à vingt essais : à chaque essai (ou, dans une variante expérimentale, une fois sur quatre) , le sujet se voit indiquer le chiffre qu'il aurait dû répondre. Ainsi on peut remplir pour chaque sujet un tableau $S \times R$: $k(s, r)$ étant le nombre de fois qu'à la couleur s le sujet a répondu en appuyant sur le bouton r . En début de série, les erreurs sont nombreuses ; elles se raréfient ensuite, de sorte que finalement la diagonale du tableau, où sont comptées les réponses exactes, contient les nombres les plus élevés. S'il n'y avait pas du tout d'erreur, l'analyse factorielle ne pourrait rien extraire du tableau $k(s, r)$: les profils des lignes, comme ceux des colonnes, seraient étrangers entre eux (chacun ne comprenant qu'un seul chiffre non nul) ; on ne pourrait par leurs proximités mutuelles établir une structure. Si, au contraire les sujets ignoraient leur tâche, tous les $k(s, r)$ seraient aux fluctuations aléatoires près, égaux entre eux : les profils seraient alors tous également plats, et, ici encore, leur comparaison ne révélerait rien. On peut donc s'attendre à ce que l'analyse soit la plus fructueuse lorsque le taux d'erreurs est modéré.

Or, comme nous le signalons en introduction, ces erreurs, (à supposer, seul cas intéressant pour nous, qu'elles reflètent la disposition de l'expérience, et non quelque cause extérieure ignorée de nous), peuvent être dues soit aux similitudes entre stimuli, soit aux similitudes entre réponses : une réponse fautive peut être soit un bouton voisin de la réponse exacte, soit la réponse qu'il faut donner à une couleur voisine de celle présentée. D'où interférence de deux structures : d'une part l'ordre des couleurs sur le spectre (puisqu'on a pris garde d'utiliser des couleurs pures, et non des teintes rabattues de noir ou lavées de blanc, et qu'on a approximativement égalisé les éclaircissements, il ne reste aux stimuli que la seule dimension spectrale ; cf. e.g. notre cours sur la vision des couleurs, 1967) ; d'autre part la disposition géométrique des boutons sur le clavier. Nous reporterons ailleurs des analyses mettant en évidence la configuration d'un clavier. Le psychologue s'intéressait ici à la généralisation sur les couleurs. Afin d'isoler ce phénomène, il donna à chaque sujet un code différent de correspondance entre couleurs et boutons ; (pour choisir ces codes, à défaut de pouvoir prendre toutes les $8!$ permutations possibles de huit objets, on recourut aux règles combinatoires de ce que les statisticiens anglo-saxons appellent "randomisation"). Ainsi pour l'ensemble des sujets, il est possible de construire un tableau de confusion k_{SS} d'où les effets de clavier sont en moyenne éliminés : on pose $k(s, s')$ égal au nombre de fois que (dans l'ensemble d'essais réalisés avec divers sujets ayant chacun leur code) la couleur s reçut pour réponse le bouton qui eût été la réponse exacte à la couleur s' . Comme nous l'avons annoncé, il ne s'agit donc pas de confusions directes entre s et s' ; mais d'erreurs reliant s à s' , et dont la seule cause jouant systématiquement dans le même sens pour tous les sujets est la proximité de s et s' sur le spectre. Il eût été possible d'utiliser les mêmes données pour construire un tableau k_{RR} : $k(r, r')$ étant le nombre de fois que le bouton r' fût donné pour réponse, alors que la réponse correcte était r . L'analyse de ce tableau eût sans doute révélé la disposition du clavier : mais ne disposant plus des résultats détaillés des expériences, nous n'avons pu faire cette analyse.

Les données analysées ici, forment finalement deux tableaux k_{ss} : l'un comptabilise les confusions faites en début d'apprentissage, l'autre concerne la fin (qui est définie, pour chaque sujet indépendamment, comme la phase qui commence au premier essai où il a donné au moins quatre réponses exactes sur huit). La somme $\sum \{k(s, s') \mid s' \in S\} = k(s)$, devrait, quelle que soit la couleur s , être égale (pour chaque tableau) au nombre des essais considérés : il n'en est pas tout à fait ainsi, parce que l'expérience présente quelques irrégularités...

1.2 - Avant de décrire les résultats de l'analyse factorielle, nous invitons le lecteur à considérer le tableau de données. Dans les situations unidimensionnelles, pour des ensembles de cardinal peu élevé (10 ou 20 ; ici 8) les manipulations de lignes et de colonnes suffisent, on le sait, à l'analyse. Le seul avantage du calcul automatique est alors de fournir des résultats ne comportant aucune décision arbitraire, et donc comparables d'une expérience à une autre. Dans les situations multidimensionnelles, (stimuli ou réponses dépendant de plusieurs paramètres sensibles), avec de grands tableaux le calcul automatique est la seule méthode praticable. Sur le tableau de fin d'expérience, la diagonale se distingue nettement : ceci prouve que même si l'ensemble des lignes et celui des colonnes avaient été soumis à deux permutations distinctes, l'on reconnaîtrait immédiatement quelle colonne répond à quelle ligne. De plus, comme il est naturel, à des couleurs voisines correspondent des profils voisins de lignes et de colonnes : mais, surtout au milieu du spectre, on serait embarrassé pour retrouver l'ordre naturel des couleurs.

Sur le tableau de début d'expérience, en revanche, on ne voit rien de très net.

Dans l'analyse factorielle de ce tableau 1.1, la suite des valeurs propres extraites est 0,156 ; 0,043 ; 0,027 ; 0,020 ; 0,011... Les facteurs d'ordre 2, 3 etc. ne nous paraissent pas le moins du monde interprétable. Le facteur 1, qui contient 60 % de l'inertie totale du tableau, alors que le facteur 2 n'en contient que 16 %, fournit des données une représentation assez satisfaisante. Sur la figure 1.1, on voit que points figuratifs des stimuli et des réponses se correspondent à peu près. De plus, l'on retrouve partiellement l'ordre des couleurs dans le spectre : après les deux groupes {Violet, Bleu, Bleu-Vert} et {Vert, Jaune-Vert, Jaune à l'intérieur desquels l'ordre est détruit, viennent l'Orangé puis le Rouge. Nous ne croyons pas qu'aucune méthode d'analyse trouve beaucoup plus dans le tableau 1.1 ; car, en début d'apprentissage (cf. supra), les sujets commettent une majorité d'erreurs dues non à la généralisation sur le stimulus ou la réponse, mais à l'ignorance quasi totale de la correspondance entre stimuli (couleurs) et réponse (touches de clavier, lettres etc.).

En fin d'expérience, au contraire, les confusions deviennent plus rares, mais la généralisation y joue un rôle prépondérant. Nous avons remarqué que la diagonale du tableau 1.2 se détache nettement ; comme ce tableau est, de plus à peu près symétrique, stimuli et réponses de même nom ne se peuvent distinguer à l'échelle à laquelle sont tracés nos graphiques. Sur le premier axe, l'ordre du spectre se retrouve, à la seule inversion près du Jaune et du Jaune-Vert (dont les longueurs d'ordre sont très voisines, cf. fig. 1.1) ; mais au centre et à l'extrémité violette du spectre, les couleurs sont plus

Tableau 1.1

Les matrices de confusion : les huit couleurs sont désignées approximativement par les initiales des noms : Rouge, Orange, Jaune, Jaune-Vert, Vert, Bleu-Vert, Bleu, Violet. Le tableau supérieur se rapporte au début de l'expérience, le tableau inférieur à la fin.

| R \ S | R | O | J | JV | Vr | VB | B | VI |
|-------|----|----|----|----|----|----|----|----|
| R | 67 | 43 | 11 | 11 | 10 | 14 | 10 | 6 |
| O | 25 | 31 | 29 | 9 | 17 | 11 | 15 | 12 |
| J | 19 | 17 | 25 | 30 | 24 | 27 | 29 | 24 |
| JV | 15 | 18 | 30 | 45 | 19 | 16 | 22 | 16 |
| VR | 19 | 16 | 25 | 29 | 35 | 31 | 21 | 19 |
| BV | 15 | 16 | 15 | 18 | 37 | 32 | 26 | 23 |
| B | 4 | 10 | 17 | 19 | 15 | 31 | 49 | 24 |
| VI | 11 | 17 | 26 | 17 | 14 | 15 | 8 | 43 |

| R \ S | R | O | J | JV | Vr | BV | B | VI |
|-------|-----|-----|-----|-----|-----|-----|-----|-----|
| R | 415 | 45 | 2 | 8 | 7 | 4 | 4 | 3 |
| O | 32 | 373 | 16 | 17 | 8 | 11 | 12 | 8 |
| J | 10 | 12 | 343 | 70 | 22 | 20 | 13 | 10 |
| JV | 6 | 19 | 50 | 303 | 31 | 23 | 18 | 6 |
| Vr | 6 | 12 | 23 | 36 | 305 | 71 | 29 | 8 |
| BV | 10 | 10 | 15 | 32 | 91 | 274 | 38 | 19 |
| B | 8 | 11 | 14 | 6 | 17 | 60 | 356 | 36 |
| VI | 3 | 5 | 22 | 13 | 11 | 13 | 24 | 403 |

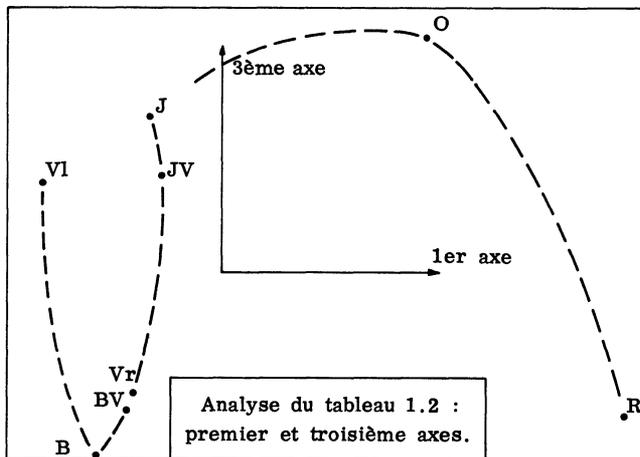
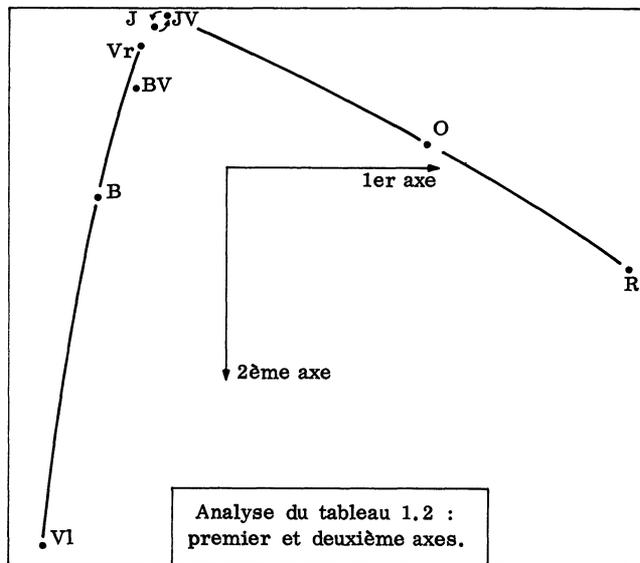
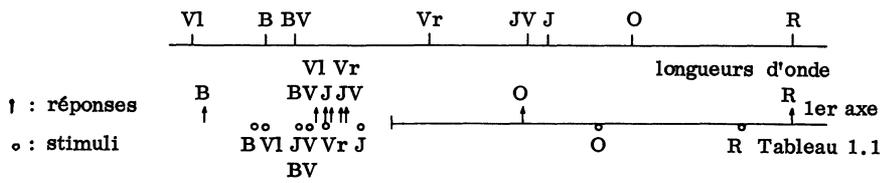


Figure 1.1 - Analyse des tableaux de confusion entre huit couleurs. On a donné aux axes la longueur unité.

étroitement resserrées qu'à l'extrémité rouge. Dans le plan des deux premiers facteurs, l'extrémité violette du spectre reçoit un étalement plus grand : le spectre forme deux tranches quasi-rectilignes qui se rejoignent au niveau du Jaune et du Jaune-Vert en formant un angle.

Cette figure est très bien connue en psychophysique de la vision des couleurs. Sans reprendre en détail une théorie classique (que nous avons exposée ailleurs : cf. Benzécri, 1967) rappelons que les stimuli lumineux agissent approximativement sur les sujets humains normaux (i.e. non daltoniens etc..) comme des vecteurs d'un cône convexe situé dans un espace à trois dimensions : la frontière de ce cône comprend une nappe courbe, ou nappe spectrale, et une facette plane (la facette des pourpres) reliant les génératrices extrêmes, (rouge et violette) de la nappe ; au centre du cône, passe la demi-droite des stimuli blancs. L'intensité lumineuse est une forme linéaire qui s'annule sur un plan (dit alychne) évidemment extérieur au cône mais auquel la nappe spectrale tend à se raccorder tangentiellement suivant la génératrice violette. L'usage est de représenter une section plane du cône des stimuli colorés : la courbe spectrale, intersection de la nappe spectrale avec le plan s'inscrit dans un triangle isocèle et rectangle ; le côté $y = 0$ est sur l'alychne ; le côté $x = 0$ est tangent à la branche bleue du spectre ; le troisième côté est tangent à la branche rouge ; enfin le blanc est au centre de gravité. C'est le diagramme de la Commission Internationale de l'Eclairage (C.I.E.) (cf. fig. 1.2).

Il est frappant que non seulement, sur le plan des deux premiers axes, les huit couleurs dessinent la courbe spectrale mais encore la direction même des axes soit celle choisie par la C.I.E. (premier axe parallèle à l'alychne ; deuxième axe faisant un angle faible avec la branche bleue du spectre). Cette rencontre n'est pas unique, comme nous le verrons plus bas (n° 1.3) en analysant d'autres données plus complètes relatives aux similitudes entre stimuli spectraux.

On sait que dans une situation unidimensionnelle, les facteurs successifs sont en général tous fonctions du premier : le deuxième facteur étant une fonction parabolique le 3ème, une cubique en S... : c'est ce qu'on peut appeler l'effet Guttman. Ici, pour autant que 8 points suffisent à définir une courbe en S, il semble que l'on soit dans ce cas. Mais ni la suite lentement décroissante des valeurs propres ($\lambda_1 = 0,758$; $\lambda_2 = 0,624$; $\lambda_3 = 0,527$; $\lambda_4 = 0,507...$), ni les graphiques (e.g. celui dans le plan 1.3) ne permettent d'affirmer absolument l'unidimensionalité observée dans le plan 1.2 : c'est trop peu répétons-le qu'un tableau 8×8 . Quant au critère du χ^2 il donne plus de cinq facteurs statistiquement significatifs ; ce qui, on le voit bien, ne signifie pas que ces facteurs (nés des conditions particulières de l'expérience) soient interprétables d'après la structure de l'ensemble des stimuli colorés.

1.3 - Après l'analyse des matrices de confusion issues de l'expérience de Richard, considérons un tableau plus ample publié par G. Eckman (1954). L'expérience d'Eckman concerne quatorze stimuli colorés quasi-monochromatiques ; ceux-ci sont présentés à l'aide d'un projecteur éclairant au travers d'un filtre une fenêtre de verre dépoli. Plus exactement, il y a deux projecteurs et deux fenêtres, ce qui permet de proposer successivement aux sujets de comparer entre eux toutes les paires de stimuli distincts. (Mais Eckman ne dit pas s'il a pris soin d'égaliser les brillances photométriques

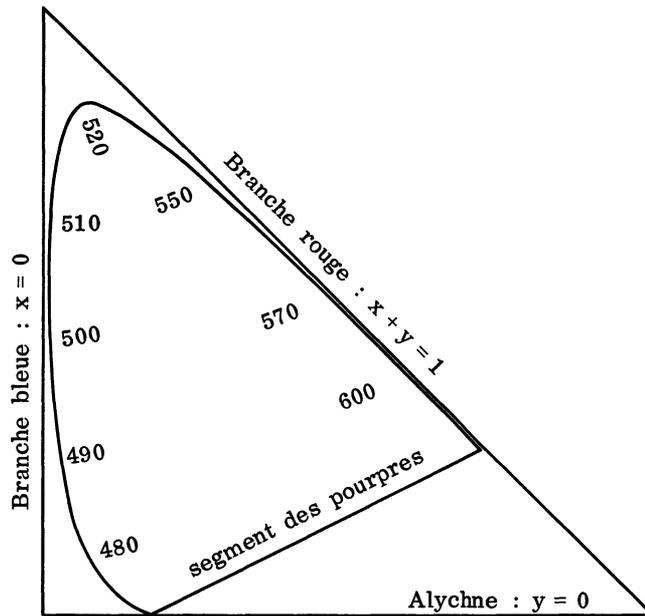


Figure 1.2 : Diagramme chromatique de la C.I.E. ; les longueurs d'onde sont mesurées en millimicrons

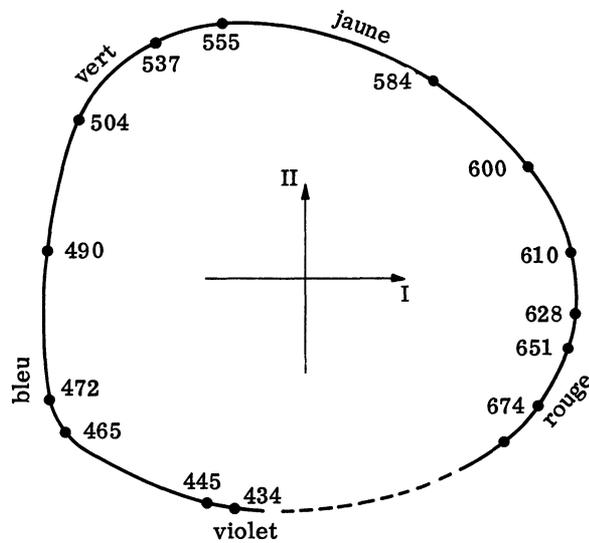
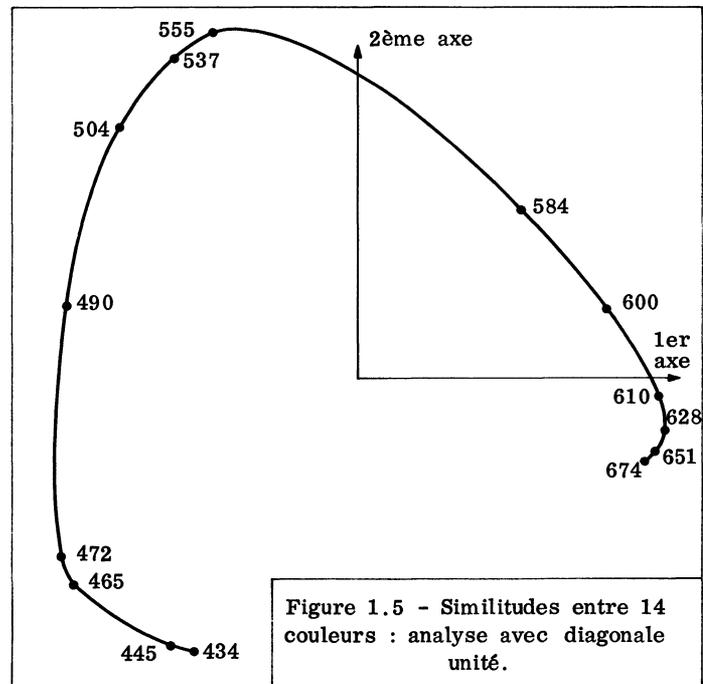
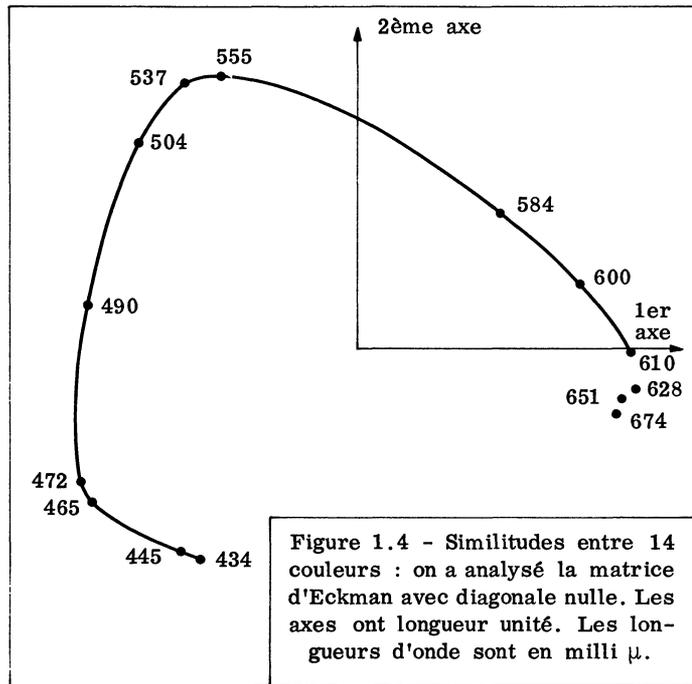


Figure 1.3 - Analyse des proximités entre 14 stimuli colorés (d'après R.N. Shepard, 1954 c)

de tous ses stimuli...) Chaque sujet évalue le degré de similitude par un nombre entier compris entre zéro et quatre. Des réponses des sujets, Eckman déduit, par calcul de moyenne, un indice de similitude pouvant varier de 0 à 1. Nous traiterons ailleurs de l'analyse de similitude. Ici nous nous bornons à donner les résultats obtenus en soumettant le tableau d'Eckman à l'analyse des correspondances, comme s'il s'agissait d'une matrice de confusion. L'article d'Eckman ne nous dit rien sur le degré de similitude d'une couleur avec elle-même : et il n'est pas sûr qu'avec la procédure expérimentale qu'il utilise on trouve pour deux stimuli identiques un indice égal à 1. A titre d'expérience nous avons analysé deux fois la matrice d'Eckman : en la complétant par une diagonale de 1, et aussi par une diagonale de zéro. On voit sur les figures 1.4 et 1.5 que les résultats obtenus dans les deux cas diffèrent peu. Comme dans la figure 1.1, les axes factoriels correspondent aux axes ($x = 0$) et ($y = 0$) du diagramme de la C.I.E. Les points (désignés sur nos figures par des longueurs d'onde mesurées en millimicrons) se rangent sur une courbe analogue à la courbe spectrale de la C.I.E. : la seule différence notable est que la branche rouge du spectre (fortes longueurs d'ondes) se termine non en droite (comme sur la courbe spectrale) mais en crochet. Avec une diagonale unité, la suite des valeurs prises est : 0,65 ; 0,39 ; 0,09 ; 0,08 ; 0,03... ; avec diagonale zéro on a : 0,55 ; 0,23 ; 0,11 ; 0,10 ; 0,09... Il est donc pleinement justifié de s'arrêter à une représentation plane ; d'autant plus que, avec diagonale unité, les facteurs 3, 4, 5 sont (cf. supra) des fonctions régulières de la longueur d'onde présentant une suite alternée de maxima et de minima ; tandis qu'avec une diagonale zéro les facteurs 3, 4, et 5 sont des facteurs inverses. Le tableau Eckman a été analysé par R.N. Shepard (1964 c) suivant sa méthode d'analyse des proximités : la représentation qu'il obtient (fig. 1.3) est satisfaisante ; mais le parallélisme avec le diagramme de la C.I.E. est moins net que dans l'analyse de correspondance. Il est intéressant de noter que, de par le principe même de l'algorithme de Shepard, l'analyse des proximités tend à diminuer les distances les plus grandes, donc à effacer les angles aigus de la périphérie des figures : cela peut-être un avantage dans certains cas (cf. infra n° 8, représentation du code Morse) ; mais ici l'aspect anguleux (quasi-triangulaire) semble dans la nature des choses, et on perd à l'atténuer.

2 - SIX STIMULI SONORES SINUSOIDAUX

L'expérience dont nous abordons maintenant l'analyse diffère peu de celle décrite au n° 1.1 : aussi pourrions-nous être plus brefs. Le psychologue (Mr^r Le Ny) utilise six stimuli sonores. Afin que n'intervienne que la seule dimension de hauteur, les sons n'ont pas de timbre : ils sont produits par des générateurs sinusoidaux. De plus afin d'éviter que le cercle de la gamme ne brouille la suite linéaire des hauteurs, les sons ont des fréquences en progression géométrique de raison 2 : c'est donc toujours la même note, mais on s'élève d'octave en octave. (On sait, cf. Shepard 1964 b, que la suite des sons sinusoidaux purs de hauteur croissante est un espace de stimuli hélicoïdal, non rectiligne : entre, par exemple le DO^3 et le DO^4 la distance pourra être moindre qu'entre le DO^3 et le FA^3 : chaque son sinusoidal pur a donc d'une part une coordonnée de hauteur, une cote Z ; d'autre part une coordonnée circulaire φ , variant e.g. de 0 à 360° quand on va du DO^3 au DO^4 . Quant aux sons complexes présentant une longue suite d'harmoniques,



de très belles expériences de R.N. Shepard ont montré qu'ils se prêtent aux effets les plus paradoxaux : il est possible de concevoir des suites finies de 12 sons qui, répétés indéfiniment, semblent s'élever sans cesse ou descendre sans cesse ; bien qu'en réalité, répétons-le, il n'y ait que douze stimuli). Des sujets apprennent à associer les six sons à six lettres : le code étant (comme plus haut avec les couleurs) différent de sujet à sujet. Il en résulte une matrice de confusion 6×6 : $k(s, s')$ étant le nombre de fois qu'un son s est associé à la lettre liée au son s' .

L'analyse de cette matrice fournit deux facteurs significatifs (au sens du critère du χ^2 , applicable ici puisqu'il s'agit d'une suite d'essais élémentaires peu corrélés entre eux), d'où une double représentation plane de l'ensemble des six sons, à la fois comme stimuli s et comme réponses s' . Les douze points se répartissent sur un arc de parabole, dans l'ordre des fréquences ; chaque réponse est proche du stimulus correspondant mais avec un décalage systématique des réponses vers les graves. L'examen du tableau des données (cf. tableau 2.1) nous assure en effet que lorsque la réponse est fautive, c'est le plus souvent que le sujet a donné une lettre qui représente un son plus grave que le stimulus qu'il a entendu (e.g. réponse R3 pour le stimulus S2). Quant à la disposition parabolique, elle provient de ce que l'on peut appeler l'effet Guttman. L'analyse fournit certes des facteurs non corrélés entre eux ; mais ceci n'exclut pas que le deuxième facteur soit une fonction quadratique du premier.

3 - VISION TACHISTOSCOPIQUE DE L'ANNEAU DE LANDOLT

3.0 - L'anneau de Landolt, cercle d'épaisseur constante interrompu en un point, est communément utilisé dans l'étude de l'acuité visuelle, notamment chez les illétrés ; le sujet examiné, à qui on présente des anneaux de plus en plus petits, doit dire la position, variable, de la lacune. Dans l'expérience ci-dessous - conçue par J.-F. Richard et H. Rouanet (1968), un anneau de 40 mm, de diamètre, 2 mm, d'épaisseur, offrant une lacune de 1 mm, est présenté en vision tachistoscopique, i.e. pendant un temps très court - $3/100$ de seconde - ne permettant pas à l'observateur de faire mouvoir un oeil pour explorer l'objet. Le sujet rend compte de la position où il a vu la lacune en marquant au crayon un point sur un cercle. Puis l'expérimentateur présente l'anneau en vision prolongée permettant au sujet de s'assurer de la position véritable de la lacune. Pour les auteurs, il ne s'agit pas d'étudier l'apprentissage de la vision tachistoscopique mais l'adaptation des sujets à diverses règles aléatoires de disposition de la lacune. En effet, le sujet, qui ne peut bien voir, accomplit plus ou moins inconsciemment une tâche de prédiction : il tend à placer la lacune soit là où elle apparaît le plus fréquemment, (effet de fréquence), soit là où elle était à l'essai précédent (effet dit de récence), etc.. Quant à nous, nous nous bornerons à rapporter ici, les analyses faites par Mme B. Le Roux, des erreurs d'observation de la lacune.

3.1 - Une première série de données analysées par B. Le Roux (cf. thèse) proviennent de 10 sujets ayant effectué chacun environ 240 essais. Dans chaque série, les positions vraies successives de la lacune sont indépendantes les unes des autres ; la loi de répartition en est uniforme sur le cercle. Plus précisément, la lacune peut occuper 40 positions équipro-

Tableau 2.1

Une expérience de Le Ny relative à un ensemble de six stimuli sonores purs

| S \ R | 1 | 2 | 3 | 4 | 5 | 6 |
|-------|-----|-----|-----|-----|-----|-----|
| 1 | 233 | 96 | 54 | 26 | 6 | 0 |
| 2 | 94 | 142 | 118 | 40 | 16 | 6 |
| 3 | 20 | 62 | 122 | 110 | 71 | 31 |
| 4 | 11 | 22 | 76 | 139 | 130 | 38 |
| 5 | 3 | 4 | 15 | 54 | 174 | 166 |
| 6 | 4 | 0 | 3 | 18 | 98 | 293 |

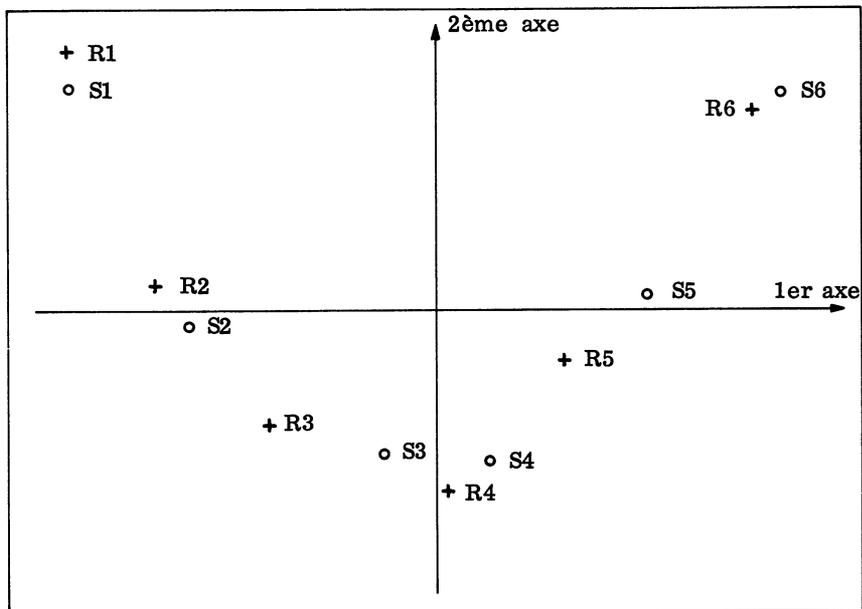


Figure 2.1 - Analyse de l'expérience de Le Ny.

bables dont les abscisses angulaires comptées à partir du point supérieur du cercle, s'échelonnent de 9 en 9 degrés à partir de 4°. Le couple stimulus-réponse (i. e. le couple position vraie de la lacune-place marquée par l'observateur) est ici un couple de points du cercle. Il n'est pas question de faire ici l'analyse sur le produit $S \times R$ de deux ensembles continus : une telle analyse intéressante dans le cas théorique où on connaît exactement la loi de probabilité du couple $(s - r)$, ne peut jamais être faite sur un ensemble fini de données. Ici, au contraire, le peu de données disponibles, même en cumulant les essais effectués par les dix sujets, suggère de réduire S et R à être des ensembles à 10 éléments, en divisant la circonférence en 10 arcs égaux découpés à partir du point supérieur, (de 0° à 36°, de 36° à 72° etc.). Chaque paire (s, r) devient ainsi une paire de nombres compris entre 1 et 10. Toutefois afin de ne pas coder trop grossièrement l'information, un point situé à proximité de la limite des deux arcs est affecté à la fois à l'un et à l'autre avec des poids convenables. Ainsi, par exemple, un couple (32°, 72°) se trouvera compté comme quatre couples, chacun de masse 1/4 : (1, 2), (2, 2), (1, 3), (2, 3). On peut présenter autrement cette règle, et en même temps la justifier, en disant que, dans le cas d'un grand nombre de couples (s, r) , elle revient à coder aléatoirement un s ou un r proche de la jonction de deux arcs, soit par le numéro de l'un de ces arcs, soit par le numéro de l'autre.

Ce codage achevé, on a une matrice de confusion 10×10 dont l'analyse factorielle fournit sur les erreurs commises par les sujets, des lois que le simple examen des données n'avait que partiellement révélées.

On retrouve d'abord, dans le plan des deux premiers facteurs, la suite des stimuli et celle des réponses rangés de 1 à 10 sur un contour à peu près circulaire. Stimuli et réponses de même rang se correspondent à peu près : toutefois il y a dans la moitié supérieure du cercle, un décalage systématique des réponses : ainsi le point R9 est entre S9 et S8 ; ce qui indique qu'une réponse fautive sera plutôt du type R9 pour S8 (ou, symétriquement par rapport à la verticale, R2 pour S3), i. e. décalée vers le point supérieur du cercle. Dans la moitié inférieure du cercle, stimuli et réponses sont le plus étroitement resserrés. En sorte que, dans l'ensemble, l'analyse factorielle fournit de la configuration circulaire étudiée une carte à échelle variable : l'échelle étant maxima au point supérieur (où les sujets commettent le moins d'erreurs) et minima au point inférieur. Les psychologues avaient bien vu sans analyse de correspondance, l'erreur systématique de décalage, vers le sommet commise dans le demi-cercle supérieur ; le grand nombre d'erreurs affectant les stimuli inférieurs ne leur avait pas échappé ; mais ils n'avaient pas encore conçu l'existence d'une carte quasi-régulière.

3.2 - Ceux qui ont étudié la géométrie différentielle n'ont pas de peine à se représenter comment un changement d'échelle variable d'un point à un autre, peut modifier la forme d'un objet : il suffit d'évoquer les ds^2 riemanniens, les représentations conformes... Pour les moins géomètres disons que c'est une sorte de déformation perspective, comme dans un portrait pris de très près où les mains sont gigantesques et l'extrémité du doigt le plus proche de l'objectif peut dissimuler un bras... Dans l'exemple analysé ici le changement variable d'échelle ne conduit à rien de tel, parce qu'un cercle, allongé dans une de ses parties, raccourci ailleurs, reste une courbe simple

formée. Mais des déformations fort grotesques se sont déjà rencontrées en psychologie, plus précisément en neurophysiologie.

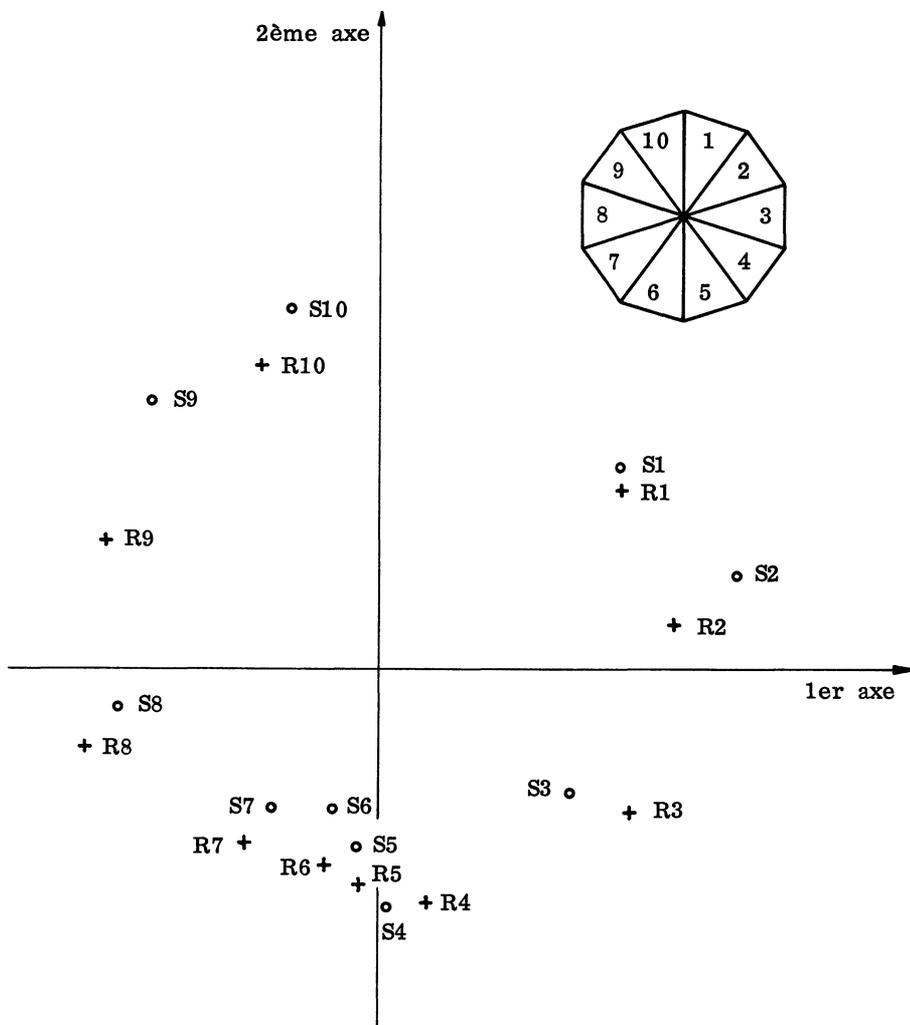


Figure 3.1 - Analyse d'une expérience de vision tachistoscopique ; on a représenté sur un polygone les numéros attribués aux 10 secteurs en lesquels on a divisé la circonférence.

On sait que la sensibilité cutanée est très variable à la surface du corps : ainsi, en appliquant un compas à pointe sèche, (l'esthésiomètre de Weber), sur le dos de la main, on constate que le sujet ne distingue deux piqûres que lorsque l'écartement des pointes est de l'ordre de 16 mm. (il faut pour mesurer ce seuil une procédure expérimentale compliquée, mêlant, e.g., aux piqûres doubles de l'esthésiomètre des piqûres faites avec une seule pointe : et finalement ce qu'on mesure est l'écartement des pointes à partir duquel le sujet a une chance sur deux de reconnaître si on le pique avec un compas ou une pointe unique) ; à l'extrémité de la langue, au contraire, le seuil esthésiométrique est de 1 mm. Si l'on rapporte toute la surface du corps à une échelle variable partout égale au seuil esthésiométrique,

(autrement dit si l'on compte les distances superficielles entre deux points non en un ou en mm., mais en nombre de seuils) et qu'on tente de faire une carte, on obtient une étrange gargouille aux lèvres énormes, à la langue proéminente, et dont les extrémités digitales, plus épaisses que les bras rappellent le portrait mal pris que nous évoquions tout à l'heure. Or cette carte n'est pas seulement un divertissement de géomètre : elle se trouve matérialisée sur les aires de projection corticale. En effet si la sensibilité cutanée varie, c'est que varie la densité des récepteurs sensoriels. De ces récepteurs sont issues des fibres nerveuses qui après des circuits compliqués se trouvent représentées dans l'écorce (la périphérie) du cerveau. Aux fibres issues de récepteurs voisins correspondent des points contigus du cortex ; à ce niveau, se matérialise dans une carte dont l'échelle (certes très réduite) est naturellement proportionnelle à l'écartement des récepteurs (à deux récepteurs cutanés limitrophes, correspondent des points du cortex dont la distance ne dépend pas du territoire cutané d'origine) donc, on l'a dit, au seuil. Cette carte cervicale, que la neurophysiologie a pu quelque peu explorer, est ce qu'on appelle l'homonculus, le petit bonhomme. Plus généralement, les psychophysiologues parlent d'homonculus chaque fois qu'à un niveau des relais nerveux se matérialise une carte (à échelle variable) d'un domaine d'objets. Et nous pouvons encore convenir d'appeler homonculus une représentation spatiale que l'analyse factorielle révèle être sous-jacente, en quelque manière, à notre conception d'un champs du réel, concret ou même abstrait. Comme le montre brillamment R.N. Shepard, à l'attention que nous portons aux objets, correspond une structure métrique, éventuellement non euclidienne. (cf. Attention and the Metric Structure of the Stimulus Space, 1964a ; article dont le titre significatif vaut d'être cité en entier).

3.3 - Revenons aux travaux de Richard et Rouanet pour exposer les analyses faites par B. Le Roux de deux autres séries d'expérience. Dans ces expériences, à la différence de celles analysées aux n° 3.1, la loi de distribution du stimulus, -c'est-à-dire de la lacune dans l'anneau de Landolt- n'est pas uniforme. La probabilité, maxima en un point du cercle que nous appellerons le mode de l'expérience, décroît régulièrement pour s'annuler au point diamétralement opposé. Dans l'une des séries d'expérience ; la courbe représentative de la densité de probabilité du stimulus en fonction de l'abscisse angulaire comprend deux segments de droite, formant un triangle avec l'axe des abscisses ; dans l'autre série la courbe est sinusoidale. Nous parlerons, en bref, de loi triangulaire et de loi sinusoidale. Notons $|\theta - \varphi|$ la mesure du plus petit des deux arcs de circonférence reliant deux points d'abscisse θ et φ ; notons θ_m l'abscisse du mode d'une distribution ; et θ'_m l'abscisse du point diamétralement opposé à θ_m ; alors on a les formules :

$$\text{Loi triangulaire : } (1/\pi^2) |\theta - \theta'_m| d\theta$$

$$\text{Loi sinusoidale : } (1/2\pi) (1 + \cos (\theta - \theta_m)) d\theta ?$$

Chaque série comprend 10 expériences, dont les modes sont les sommets d'un décagone régulier ; et chaque expérience, faite avec un sujet différent, comprend environ 300 essais.

Bien que stimulus et réponse puissent occuper sur le cercle un grand nombre de positions, on construit, comme ci-dessus (cf. n° 3.1), une ma-

trice de confusion réduite : la seule différence est que l'on a divisé la conférence en 8 arcs et non plus en 10. En groupant les couples stimulus-réponse de chacune des deux séries, on remplit deux matrices de confusion 8×8 : l'une correspondant à la loi triangulaire, l'autre à la loi sinusoidale. (L'analyse d'une seule expérience isolée serait intéressante ; mais il y aurait trop peu de données pour permettre de calculer des facteurs stables).

Dans l'une et l'autre série, l'analyse des correspondances permet de retrouver, comme au n° 3.1, l'ordre circulaire des stimuli et des réponses (cf. fig. 3.2 ; 3.3). Avec, la loi triangulaire, l'effet d'échelle variable, que nous avons rapproché de l'homoculus des neurophysiologues, disparaît quasi-totalement : les points 1 à 8 sont à peu près équidistants sur un cercle ; toutefois le premier axe qui porte 10 % de plus d'inertie que le second, demeure à peu près exactement horizontal, ce qui prouve que la verticale garde une place distinguée. Il est vraisemblable que, pour chaque sujet, le mode de la loi des stimuli qu'on lui présente tend à prendre dans son attention, la place d'abord dévolue au point supérieur du cercle ; d'où il résulte que pour l'ensemble de la série, ce que Shepard appellerait la métrique d'attention est, en moyenne, quasi-uniforme sur le cercle. Avec la loi sinusoidale, le rôle privilégié du point supérieur se fait encore sentir sur la moyenne de la série ; les points de la moitié inférieure du cercle, bien que nettement distincts entre eux, sont plus resserrés que ceux de la moitié supérieure.

3.4 - Notons pour conclure qu'afin de mettre en évidence la nouvelle échelle d'attention induite sur le cercle par la loi des stimuli, il conviendrait de rassembler autrement en une matrice les résultats de chaque série d'expérience : chaque exemple $s - r$, obtenu dans une expérience de mode Θ_m , serait compté comme un couple $(s - \Theta_m, r - \Theta_m)$; autrement dit serait soumis à la rotation amenant Θ_m au point supérieur du cercle. En cumulant ainsi les résultats fournis par 10 sujets avec 10 modes distincts, on pourrait, avec une précision satisfaisante, dessiner un homoculus dans le système d'axe lié à l'expérience.

Les recherches suggérées dans ce n° 3-4 ont été faites par Madame B. Le Roux qui les a publiées dans sa thèse ; Isup 1970 - Les résultats analogues à ceux décrits ici, sont fort satisfaisants).

4 - HUIT STIMULI DISPOSES EN OCTOGONE

L'expérience que nous rapportons maintenant a été conçue par R.N. Shepard, pour étudier ce que devient, dans le comportement des sujets, la forme que revêt l'ensemble des stimuli dans un espace physique rapporté à des paramètres de description. A cette fin on utilise huit stimuli, qui sont cercles dont un rayon est marqué. Représentés dans un plan avec pour abscisse le diamètre du cercle, pour ordonnée l'angle d'inclinaison du rayon, ces huit stimuli sont les sommets d'un octogone régulier (cf. fig. 4.1) (les deux axes sont gradués de telle sorte qu'à des déplacements égaux sur l'un et l'autre correspondent des variations également discriminables des stimuli ; autrement dit, cf. supra, le seuil de discriminabilité est pris pour unité sur chaque axe). D'une expérience d'apprentissage d'association entre ces stimuli et des lettres, il résulte (comme dans les expériences 1 et 2 ci-dessus) une matrice de confusion 8×8 . Renvoyant au mémoire original (Shepard 1964a) pour une discussion fine de l'expérience, nous nous bornons à donner

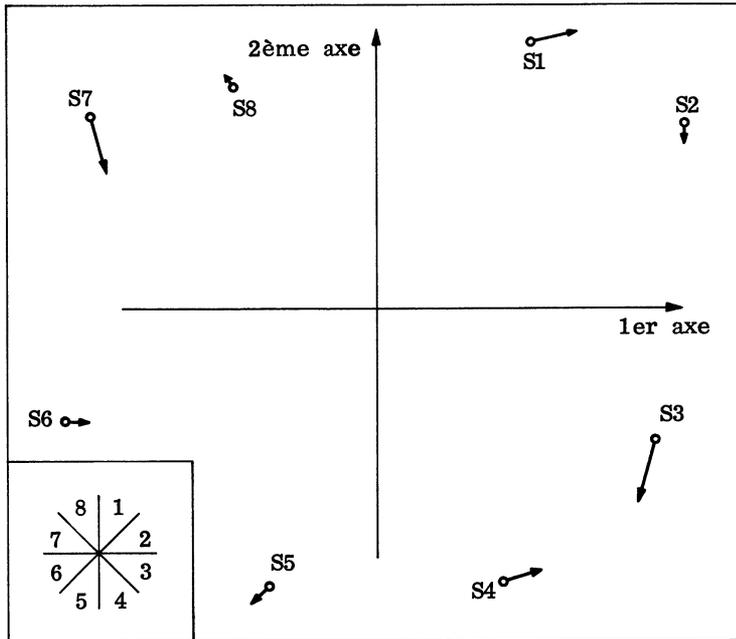


Figure 3.2 - loi triangulaire des stimuli

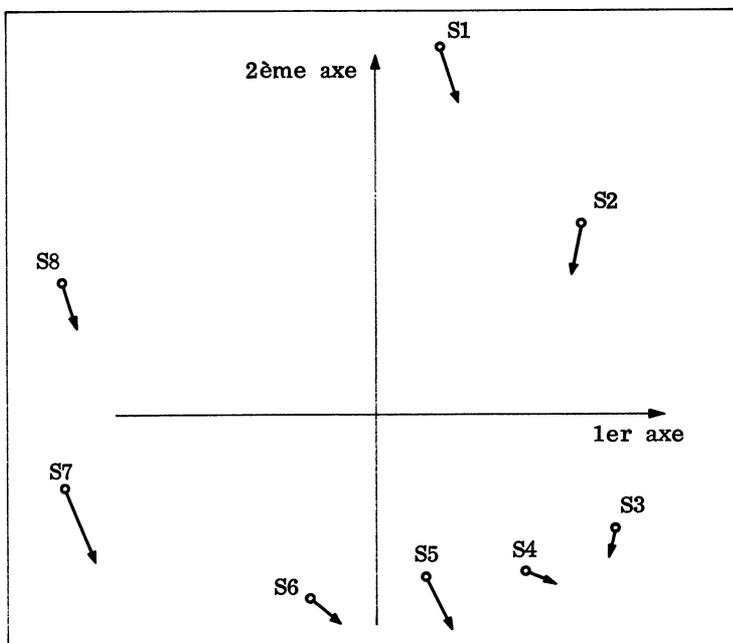
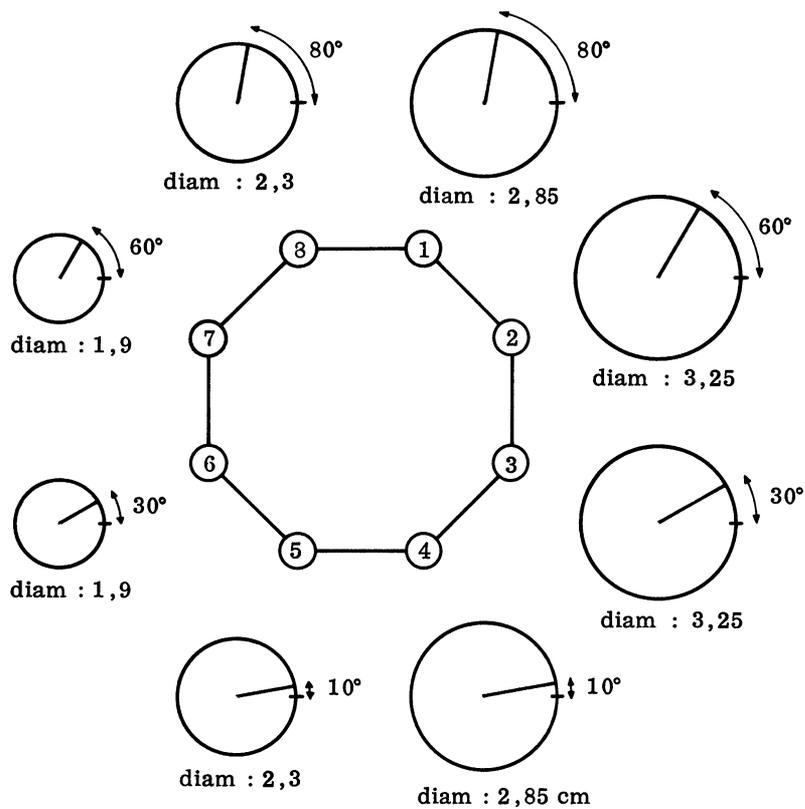


Figure 3.3 - loi sinusoidale des stimuli



Les stimuli de l'expérience de Shepard

Figure 4.1

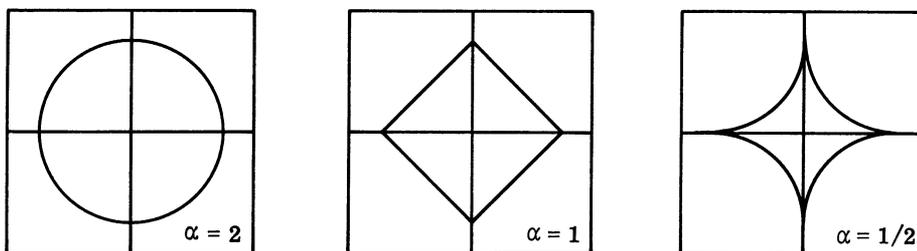


Figure 4.2 - Trois exemples de courbes d'équation $|x|^\alpha + |y|^\alpha = 1$

Tableau 4.1

La matrice de confusion des huit stimuli en octogone

| S \ R | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 1 |
|-------|-----|-----|-----|-----|-----|-----|-----|-----|
| 2 | 175 | 70 | 32 | 15 | 28 | 37 | 24 | 18 |
| 3 | 71 | 230 | 41 | 20 | 48 | 34 | 14 | 17 |
| 4 | 25 | 80 | 263 | 67 | 34 | 23 | 15 | 15 |
| 5 | 15 | 39 | 89 | 247 | 49 | 38 | 14 | 16 |
| 6 | 29 | 31 | 21 | 45 | 225 | 92 | 25 | 10 |
| 7 | 61 | 21 | 17 | 26 | 75 | 184 | 53 | 18 |
| 8 | 13 | 5 | 28 | 18 | 21 | 30 | 215 | 102 |
| 1 | 23 | 12 | 34 | 5 | 10 | 10 | 80 | 207 |

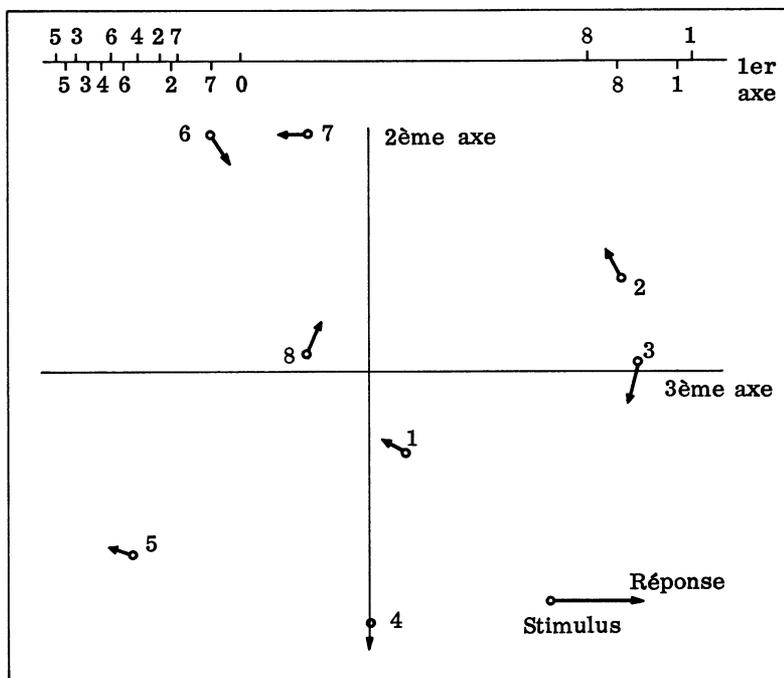


Figure 4.3 - Analyse factorielle de l'expérience faite avec ces stimuli

ici les résultats (déjà publiés cf. Benzécri 1966a) de l'analyse de la matrice de confusion.

On a pu extraire trois facteurs significatifs, d'où un diagramme spatial (tridimensionnel). L'importance relative des facteurs étant 2, 1, 1, il est légitime de représenter le premier seul sur un axe et les deux suivants sur un plan. Le premier facteur sépare nettement (cf. fig. 4.2) des six autres stimuli la paire (1,8). Dans le plan des deux facteurs suivants, 2, 3, 4, 5, 6, 7 apparaissent en cercle dans l'ordre attendu ; 1 et 8 se projettent à peu près au centre de la figure. Entre les positions des stimuli et réponses de même rang il ne semble y avoir aucune différence systématique significative. Que la paire (1,8) s'oppose aux six autres sommets sur le premier facteur s'explique sur la matrice de confusion parce que 1 et 8, souvent confondus entre eux sont peu confondus avec les autres stimuli, même avec leurs voisins immédiats 2 et 7 : si on range lignes et colonnes dans l'ordre 2, 3, 4, 5, 6, 7, 8, 1, on voit apparaître le long de la diagonale deux blocs carrés successifs, l'un 6×6 , l'autre 2×2 , en dehors desquels il n'y a que des nombres petits ou très petits. On sait que si les deux rectangles 6×2 , 2×6 ne contenaient que des zéros, le premier facteur donnerait exactement une partition, étant constant d'une part sur (1,8) (S ou R) et d'autre part sur (2, 3, 4, 5, 6).

C'est donc cette structure que l'on retrouve ici approximativement. Quant à l'interprétation psychologique, nous avons noté précédemment (1966) ; la place de la paire (1,8) suggère que pour les sujets la dimension angulaire, marquée des deux directions privilégiées horizontale et verticale, n'est pas homogène. A la lumière de l'expérience de Richard et Rouanet analysée ci-dessus (n° 3), il nous paraît maintenant vraisemblable que l'attention des sujets se portant principalement sur la partie supérieure du cercle, ceux-ci mettent à part (sans les distinguer particulièrement entre eux par leur diamètre) les stimuli 1 et 8 dont les rayons marqués, inclinés à 80° , se terminent presque au point le plus haut du cercle. Quant au plan des facteurs 2 et 3, on y notera que, comme l'a signalé Shepard les paires de stimuli (1,8), (2,3), (4,5), (6,7), reliés par un côté de l'octogone parallèle à un axe (i.e. ne différant que suivant une dimension) sont plus proches que (1,2), (3,4), (5,6) ou (7,8) (qui diffèrent sur les deux axes). Ceci suggère que entre deux stimuli reliés par le vecteur (x, y) , la distance psychologique est supérieure à la valeur $(x^2 + y^2)^{1/2}$ que donne le théorème de Pythagore. L'ensemble des stimuli (x, y) situés à une distance donnée σ (ou seuil d'un centre (x_0, y_0)) ne serait donc pas un cercle (C_2), mais plutôt une figure (C_α), avec $\alpha < 2$ (pour $\alpha = 1$ on a un carré, et pour $\alpha < 1$ on a une étoile non convexe, cf. fig. 4.3).

$$C_\alpha = \{(x, y) \mid [|x - x_0|^\alpha + |y - y_0|^\alpha]^{1/\alpha} = \sigma\}$$

Comme le dit Shepard, cette formule pour $\alpha \neq 2$ n'est pas conciliable avec une géométrie localement euclidienne de l'espace perceptif. Il est particulièrement curieux que comme exemple d'espace de stimuli à peu près localement euclidien, Shepard cite le corps chromatique (continuum tridimensionnel des stimuli colorés, variant à la fois en dominante spectrale, pureté et intensité lumineuse ; cf. supra n° 1.2), alors que des expériences récentes (Parra 1966, cité dans notre cours, 1967) ont révélé que le lieu des stimuli situés à une distance donnée d'un centre est une courbe étoilée du type $C_{1/2}$. Ainsi, la structure neurophysiologique très complexe de la per-

ception des stimuli multidimensionnels, se montre quelque peu dans la structure géométrique accessible à l'expérience.

5 - UN JEU D'INSTAURATION D'UN CODE

L'objet ultime de l'expérience dont nous analysons maintenant les résultats est d'étudier comment, dans l'instauration d'une langue, un lien s'établit entre d'une part un domaine d'expression, (e.g. l'ensemble des sons que peut produire l'appareil vocal humain) et d'autre part un domaine d'objets à désigner, domaine certes structuré mais dont l'ordre global échappe à la conscience du sujet parlant ; comment des signaux isolés (ou plutôt des domaines disjoints de signaux, puisque les diverses émissions d'un même signal se distribuent autour d'une moyenne), les mots, en viennent à désigner des classes d'objets qui recouvrent l'ensemble des objets à désigner ; comment, entre deux ensembles, ainsi qu'une araignée tend sa toile les sujets, en communiquant, tissent la langue. De cet ambitieux programme (cf. Benzécri 1965 ; 1966b) nous parlerons peu, nous bornant à expliquer l'origine de la matrice de confusion analysée ici. Disons donc comment se déroule l'expérience, faite en collaboration avec J.F. Richard.

L'ensemble A des objets à désigner est une carte d'Europe, où figure vingt-cinq pays : les objets. L'ensemble B des signaux est un segment de droite d'une dizaine de centimètres ; les signaux sont les points de ce segment. (La même expérience, à des détails près peut être faite avec d'autres ensembles...). Les sujets sont par groupe de deux, se tournant le dos ; chacun a devant lui un exemplaire de l'ensemble A, (une carte d'Europe). Un expérimentateur surveille chaque groupe ; il détient de nombreux exemplaires vierges de l'espace B des signaux (je veux dire du papier réglé !). Avant de commencer l'expérience, l'expérimentateur donne les consignes aux sujets :

"Vous allez participer à une sorte de jeu. Je vais montrer à l'un de vous un pays sur la carte : vous devez essayer d'indiquer à votre camarade ce pays, mais pour cela vous aurez seulement le droit de marquer un point sur une ligne horizontale. En voyant ce point, votre camarade devra essayer de deviner quel est le pays que vous avez voulu lui désigner. Au début, vous opérerez un peu au hasard, mais petit à petit, vous arriverez à vous comprendre en ayant comme seul moyen d'expression, la possibilité de marquer un point sur une ligne. A celui d'entre vous qui aura essayé de deviner la première fois, je montrerai ensuite un pays sur la carte, et il devra le désigner par un point pour l'indiquer à son camarade qui aura à deviner à son tour, et ainsi de suite. D'autres groupes sont en concurrence avec le vôtre. Selon que vous aurez deviné un pays proche ou lointain de celui qui était désigné, votre groupe gagnera des points ou en perdra. Vous devez essayer de gagner le plus de points possibles".

Ainsi, dans chaque groupe, chacun des deux sujets demande à tour de rôle un pays (choisi au hasard dans A par l'expérimentateur) en plaçant un point sur un segment ; l'autre écrit en regard, (à gauche) le pays qu'il croit lui être demandé ; enfin le demandeur écrit (à droite) le pays qu'il voulait désigner :

| | | |
|---------|--------|---------|
| réponse | signal | demande |
| France | | Italie |

L'expérience terminée, on réunit les feuilles de messages en un long rouleau, dont chaque ligne est un segment portant un point avec, dans les marges les interprétations de ce point à l'émission (à droite) et à la réception (à gauche).

Les données analysées ici proviennent de huit couples de sujets. Pour la plupart, en quelques dizaines d'essais, le code suivant (certes imparfaitement efficace, surtout à l'est de l'Europe) s'établit : les pays, rangés d'Ouest en Est suivant la longitude sont codés de gauche à droite sur le segment. La Russie très tôt placée par presque tous à l'extrémité droite, n'est l'objet que de peu d'erreurs. Réunissant tous les résultats d'expérience, on calcule un code moyen (cf. fig. 5.1) pour les demandes et les réponses : e.g. le code moyen de la France, en demande (resp. réponse), est l'abscisse moyenne des points signaux marqués sur des lignes à droite (resp. gauche) de laquelle il est écrit France. Et l'on construit une matrice de confusion 25×25 , $k(s, s')$ étant le nombre de lignes messages comportant à droite la demande s et à gauche la réponse s' .

Le premier facteur extrait (cf. fig. 5.2) oppose la Russie à tous les autres pays. De même que dans l'expérience de Shepard (cf. n° 4) les huit stimuli se scindaient en deux classes (1,8) et (2, 3, 4, 5, 6, 7), on a ici une partition en deux classes dont l'une réduite à la seule Russie. Et l'explication est que la Russie est très rarement confondue avec les 24 autres pays : approximativement, la matrice est faite de deux carrés 24×24 et 1×1 , en dehors desquels il n'y a que des zéros.

Sur le deuxième axe factoriel (cf. fig. 5.3), La Russie est à l'origine. Les vingt quatre autres pays se rangent dans l'ordre de leur longitude. Cela était possible, vu les codes des sujets. Mais, fait lui très remarquable, sur l'ensemble des 24 pays le code moyen et le 2ème facteur se corrèlent à 96. Ainsi l'analyse d'un tableau de nombres entiers (matrice de confusion) révèle une dimension continue mesurable. Ce succès, (qui n'est pas unique, cf. infra le code Morse) laisse espérer que l'analyse factorielle puisse révéler, de façon sûre certains maillons de l'activité mentale ; passer de dénombrements de faits observables à la mesure de dimensions cachées.

6 - SEIZE PHONEMES CONSONANTIQUES DE LA LANGUE ANGLAISE

L'étude des confusions entre sons d'une langue naturelle, intéresse également le linguiste et l'ingénieur. Mais le domaine à explorer est très vaste : et, à supposer qu'on disposât de ressources illimitées, il resterait l'embaras de choisir quelles expériences faire. Aussi les données sont-elles rares : et l'on ne s'étonnera pas que les beaux résultats publiés en 1955 par Miller et Nicely, continuent, bien qu'incomplets, d'intéresser vivement les statisticiens (cf. e.g. S. Johnson 1967). Avant d'en rapporter l'analyse, décrivons brièvement ces données.

6.1 - Cinq personnes (de sexe féminin) participaient aux expériences ; dans chaque série l'une d'elles dictait à ses quatre compagnes une suite aléatoire de syllabes séparées par des silences d'environ 2 sc. Toutes les syllabes étaient formées d'une consonne initiale suivie de la voyelle |a| (comme dans father) ; la consonne était choisie de façon équiprobable parmi les 16 phonèmes suivants :

Demandes : C^D

Réponses : C^R

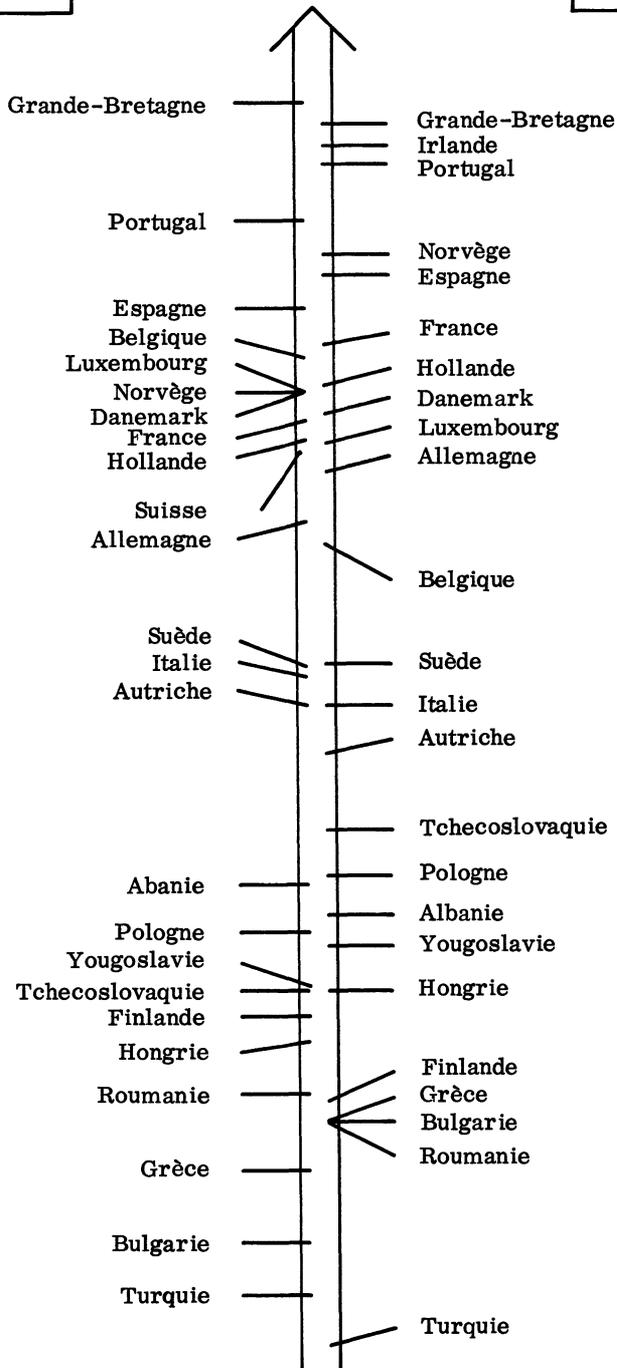


Figure 5.1 - Le code moyen

Demandes : f_1^D

Reponses : f_1^R

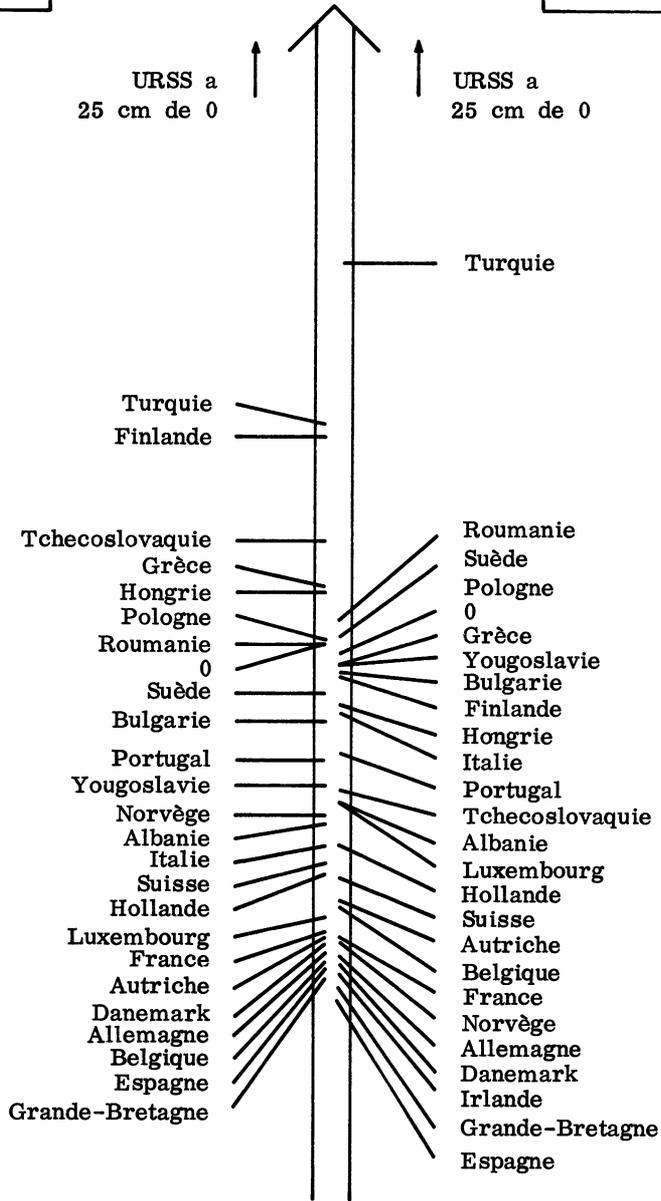


Figure 5.2 - Instauration de code ; premier facteur.

Demandes : f_2^D

Réponses : f_2^R

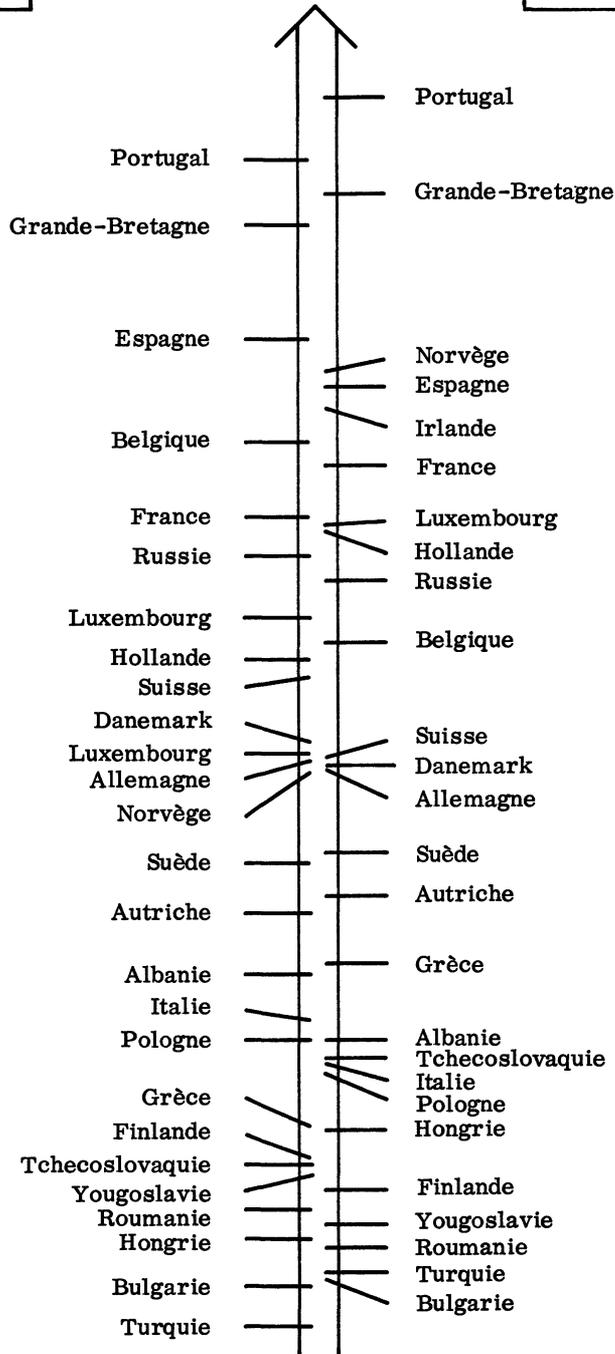


Figure 5.3

$|p|, |t|, |k|, |f|, |\theta|, |s|, |\int|, |b|, |d|, |g|, |v|, |\text{ʃ}|, |z|, |3|, |m|, |n|.$

(Pour les lecteurs non accoutûmés aux notations des phonéticiens, précisons qu'on note : $|\theta|$, le th dur de thing ; $|\text{ʃ}|$ le th doux de the ; $|\int|$, (signe d'intégrale) le sh de she correspondant au ch français ; $|3|$ un z cursif ou un chiffre 3 décalé) le s de measure, correspondant au j français). Comme la dictée se faisait par microphone et casque téléphonique, on pouvait de série en série modifier les conditions d'écoute (bruit et bande passante); nous reviendrons sur ce point très important. Les données publiées par Miller et Niceley remplissent 17 matrices de confusion 16×16 : chaque matrice correspond à une condition d'écoute particulière et on y a cumulé les données de vingt suites écrites par chacune des personnes sous la dictée des quatre autres...

Tout en se félicitant de disposer d'un parail matériau fruit d'un travail monumental, il faut souligner que les données ne concernent que 16 consonnes, alors qu'il y a en anglais une trentaine de phonèmes consonantiques (en y comprenant les consonnes doubles) sans parler des nombreuses voyelles et diphtongues et des semi-consonnes. D'autre part, attribuer aux consonnes utilisées la même probabilité ($1/16$), au lieu de probabilités proportionnelles à leurs fréquences dans la langue anglaise n'est pas sans inconvénient. Considérons par exemple les deux phonèmes d et g souvent écrits l'un pour l'autre. Sur l'ensemble des tableaux de Miller & Niceley, on trouve que d a été écrit près de mille fois à la place de g ; tandis que le nombre de cas où les sujets ont entendu g alors qu'on lisait d est deux fois moindre. Comment expliquer cette dissymétrie ? Est-ce que dans les conditions pénibles d'écoute réalisées par les expérimentateurs le g est brouillé de sorte qu'il en vient à ressembler à un d plus souvent que l'inverse ne se produit ? Cette explication ne peut être écartée sans un minutieux examen. Mais supposons plus simplement que dans 1500 cas environ (pour ne rien dire de ceux où la réponse exacte a été donnée par hasard) les effets cumulés du brouillage et de la fatigue aient réduit le son émis, qu'il soit un g ou un d, à pouvoir être aussi bien perçu comme g que comme d : cet effet de fréquence est bien connu, et nous l'avons évoqué plus haut à propos de l'expérience de Richard et Rouanet sur l'anneau de Landolt (n° 3). Or dans la langue anglaise, la fréquence du $|d|$ est environ six fois celle de $|g|$; tandis que dans les dictées les deux phonèmes ont la même fréquence : on peut donc, a priori, expliquer le rapport 2 trouvé ici, comme un compromis entre 6 et 1.

L'idéal serait-il, afin de rendre au tableau de confusion sa symétrie d'expérimenter en conservant aux phonèmes consonantiques les fréquences qu'ils ont dans l'anglais parlé ? Il faudrait encore varier les voyelles, car à l'approche de la voyelle $|a|$, les consonnes sont vraisemblablement ouvertes de sorte que leurs similitudes mutuelles ne sont pas ce quelles sont en moyenne. On devrait donc, même si a priori on ne s'intéresse qu'aux consonnes, expérimenter au moins sur l'ensemble des syllabes formées d'une consonne suivie d'une voyelle, en conservant aux syllabes les fréquences propres à l'anglais. C'est là une tâche d'autant plus immense que certaines de ces syllabes sont très rares. Or considérons les consonnes, voisines l'une de l'autre, v et ʃ ; plaçons après elles le son (sorte de e, comme dans le français je) que les phonéticiens représentent par un e retourné : avec v on a une syllabe rare : avec ʃ on a l'article défini the. Ainsi effets de fréquence re-

joignent effets de signification... Il est temps de clore notre critique en approuvant Miller et Nicely d'avoir tranché dans le vif du sujet pour s'assigner une tâche réalisable !

Définissons maintenant les 17 conditions d'écoute dans lesquelles opéraient les psychologues. Il faut pour cela des termes techniques, mais ceux-ci sont aujourd'hui fort usités, étant connus de ceux qui manipulent les boutons des potentiomètres de tonalité... Les auteurs considèrent que la réponse en fréquence du système d'amplificateur et d'écoute est constante de 200 à 6500 cps : ils utilisent pour modifier la dictée d'une part un générateur de bruit blanc (i.e. également distribué dans toutes les fréquences) limité en dessous de 7000 cps par un filtre passe-bas ; d'autre part des filtres passe-haut et passe-bas qui restreignent le message parlé à une bande plus étroite que celle de 200 à 6500 propre à la chaîne électro-acoustique. Les données des tableaux I à VI correspondent à des expériences faites avec une bande passante maxima (200, 6500), le rapport signal sur bruit croissant en progression géométrique de -18 dB à 12 dB (Rappelons que l'on exprime en décibels le logarithme d'un rapport de puissances sonores : un rapport 0 dB correspond donc à l'égalité des puissances du signal et du bruit ; pour 6 dB (resp - 6 dB) le signal est quatre fois plus (resp moins) fort que le bruit). Dans tous les autres cas, le rapport signal sur bruit est de 12 dB, mais la bande passante est réduite. On a les 17 conditions suivantes, en trois séries :

| | | | | | | | |
|----------------------|---|-------|-------|-------|-------|-------|-------|
| n° du tableau | : | I | II | III | IV | IV | IV |
| inf. de bande en cps | : | 200 | 200 | 200 | 200 | 200 | 200 |
| sup. de bande en cps | : | 6 500 | 6 500 | 6 500 | 6 500 | 6 500 | 6 500 |
| rapport en dB | : | - 18 | - 12 | - 6 | 0 | 6 | 12 |
| | | | | | | | |
| n° du tableau | : | VII | VIII | IX | X | XI | XII |
| inf. bande en cps | : | 200 | 200 | 200 | 200 | 200 | 200 |
| sup. bande en cps | : | 300 | 400 | 600 | 1 200 | 2 500 | 5 000 |
| rapport en dB | : | 12 | 12 | 12 | 12 | 12 | 12 |
| | | | | | | | |
| n° du tableau | : | XIII | XIV | XV | XVI | XVI | XVII |
| inf. bande en cps | : | 200 | 1 000 | 2 000 | 2 500 | 3 000 | 4 500 |
| sup. bande en cps | : | 5 000 | 5 000 | 5 000 | 5 000 | 5 000 | 5 000 |
| rapport en dB | : | 12 | 12 | 12 | 12 | 12 | 12 |

6.2 - On a soumis à l'analyse des correspondances les 17 tableaux 16 x 16 publiés par Miller et Nicely : ainsi que ces mêmes tableaux symétrisés (suivant la formule $k_{sym}(i, i') = (k(i, i') + k(i', i))/2$). On sait que l'analyse d'un tableau sur $I \times I$ non symétrique fournit de l'ensemble I deux représentations : chaque élément i de I pouvant être considéré comme désignant une ligne ou une colonne. Dans le cas d'un tableau symétrique ces deux représentations coïncident exactement (à moins que pour un facteur inverse l'on n'ait : $\forall i \in I : G(i) = -F(i)$). Dans la plupart des analyses de ceux issus de tableau symétrisé : les deux représentations de l'ensemble I issues du tableau k et celle issue du tableau k_{sym} ne se distinguent pratiquement pas. Il y a toutefois deux exceptions : le tableau VI, qui comporte si peu de confusions que l'analyse en est peu sûre et varie par symétrisation ; et le tableau I où le taux de confusion est si élevé (du fait du bruit qui permet à peine de distinguer les consonnes) que les deux représentations, issues

du tableau brut différent sensiblement (i.e. $F(i) \neq G(i)$) ; sans que toutefois ces différences, certes nettement visibles, n'affectent l'organisation des 16 consonnes en trois classes (cf. infra) dans le plan des deux premiers facteurs. Dans ces conditions, pour simplifier les figures, on a représenté exclusivement les résultats d'analyse des tableaux symétrisés. Quant à la validité statistique des facteurs représentés elle est toujours nettement satisfaisante selon l'épreuve du χ^2 . Sauf dans l'analyse du tableau I, où les valeurs propres sont très faibles et les facteurs 1, 2 sont à la limite permise ; nous avons pourtant représenté ces facteurs, parce qu'ils s'accordent avec les résultats de l'analyse du tableau II elle-même nettement significative.

Signalons enfin que puisqu'il est indifférent de substituer à un facteur $F(i)$ la fonction opposée $-F(i)$, on a transformé ainsi les coordonnées calculées par la machine chaque fois que cela était commode pour donner à une nouvelle figure une orientation qui permette de le mieux comparer à une figure déjà tracée.

6.3 - Débutons l'examen des résultats par le tableau VII. La bande passante est alors très étroite, les confusions sont nombreuses ; les facteurs qui les régissent se trouvent être bien connus des linguistes. Ce nous sera l'occasion d'exposer la classification généralement préconisée pour les phonèmes consonantiques. Sur le plan des deux premiers facteurs apparaissent trois îlots :

$$\{k, t, p, f, \theta, s, \int\} ; \{g, d, b, v, \partial, Z, 3\} ; \{m, n\}.$$

Ces groupes sont définis par ce que les phonologues appellent les traits pertinents de sonorité et de nasalité. On sait que les consonnes s'opposent aux voyelles comme des bruits transitoires à des sons relativement musicaux : cependant certaines consonnes comportent un élément musical. Prononçons les syllabes ba et pa en gardant les oreilles bouchées afin que la voix résonne : on remarque dans le cas du ba une vibration, un bourdonnement qui provient du larynx (des cordes vocales) et prend naissance avant que ne débute la voyelle a ; il n'en est pas ainsi avec le pa. On dit que la consonne b est sonore, tandis que p est sourde. Parmi les 16 consonnes de l'expérience, $\{k, t, p, f, \theta, s, \int\}$ sont sourdes ; les autres sont sonores. Parmi les sonores, deux sont prononcées de telle sorte que la cavité buccale communique avec la cavité nasale, où la vibration vient résonner : ce sont les nasales (m, n). Le premier facteur oppose les sourdes placées d'un côté de l'origine, aux sonores, placées de l'autre côté : comme il est naturel du fait de leur sonorité renforcée, les nasales occupent une position extrême. Le second facteur divise les sonores en nasales et non-nasales (orales), les sourdes, déjà séparées des sonores sont au voisinage du zéro. Il est tout à fait satisfaisant que le second facteur qui rend compte d'une distinction qui ne peut diviser les sourdes (pas de vibration, donc pas de résonnance nasale) soit à peu près nul pour ces consonnes.

Passons au plan des facteurs 3 et 4. Les consonnes sourdes et les sonores orales y sont les unes et les autres scindées en deux îlots. On reconnaît l'opposition entre occlusives et fricatives : e.g. entre le p et le f : p, produit comme par le choc de l'air qui ouvre brusquement les lèvres closes ; f, par le frottement d'un courant filtrant sur les lèvres entrouvertes .

Les nasales {m, n} qui sont étrangères à cette distinction, puisque le trajet nasal est en permanence ouvert à la voix, tandis que la voix buccale est interrompue, sont à l'origine dans le plan (3, 4). On peut noter avec satisfaction que sur le quatrième axe l'origine sépare exactement le groupe des occlusives (sourdes ou sonores) {p, t, k, b, d, g} de celui des fricatives ; mais le troisième facteur n'est pas interprétable. Sans trop demander aux facteurs, il faut plutôt souligner que les quatre groupes :

occlusives sourdes, occlusives sonores, fricatives sourdes, fricatives sonores, bien que définies par deux caractères dichotomiques ne forment pas un carré (ou un rectangle) mais un tétraèdre, qui s'inscrit dans un parallélépipède à base carrée (le calcul des dimensions se fait facilement d'après les facteurs : la base du parallélépipède est le plan (3ème, 4ème) ; sa hauteur est la droite du plan (1er, 2ème) joignant les sourdes aux sonores orales ; de telles épures servent parfois à l'interprétation d'une analyse factorielle...) L'opposition occlusive fricative qui se manifeste deux fois, au sein des sourdes et au sein des sonores orales, apparaît ainsi suivant deux directions différentes (et mêmes orthogonales) ; les diagonales du carré de base.

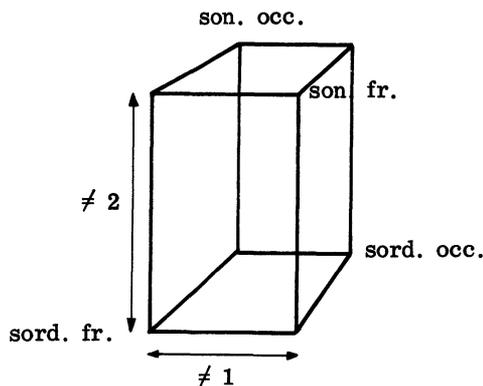
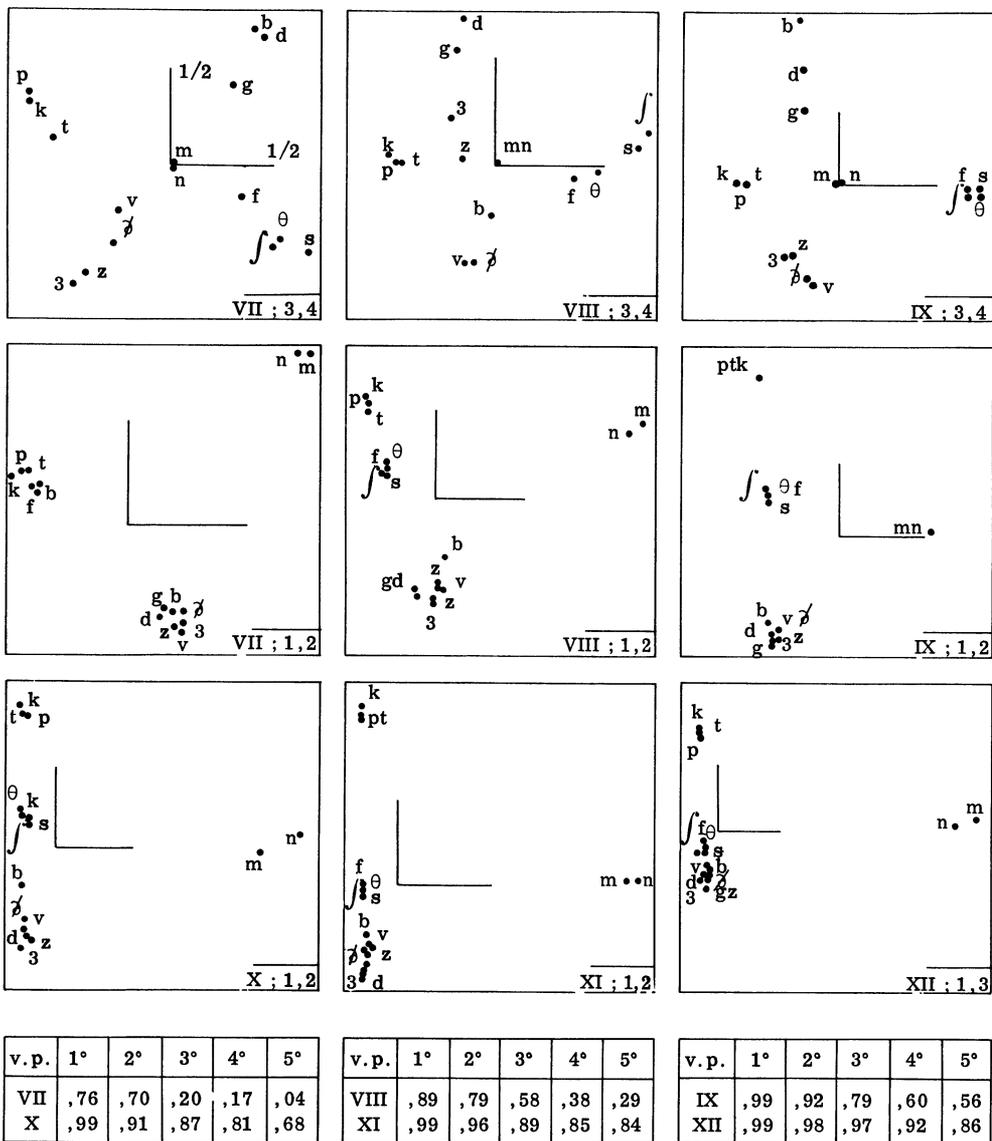


Figure 6.1 - Disposition en tétraèdre des quatre groupes de consonnes orales.

Les trois oppositions déjà décrites ne suffisent pas à distinguer les unes des autres toutes les consonnes. Ainsi {b, d, g} sont toutes trois des sonores nasales occlusives. La distinction entre ces trois sons peut se faire d'après le point d'articulation, i.e. le niveau du trajet buccal où se produit soit l'occlusion soit le frottement ; ainsi b est labiale (lèvres) d, dentale et g gutturale (gorge et plutôt arrière bouche). On notera que les dentales pour lesquelles l'ouverture de la voie nasale s'accompagne d'une occlusion buccale, ont aussi un point d'articulation. Il reste encore à distinguer θ de ϕ de Z : l'articulation de θ , ϕ n'est certes pas exactement celle de s, Z : mais pour ne pas s'aventurer dans des distinctions délicates et controversées, Miller et Nicely notent simplement que les fricatives {s, \int , Z, $_3$ } s'opposent aux douze autres consonnes étudiées en ce qu'elles sont des longues. Il resterait à voir comment les différences consignées ici, qui (la durée mise à part) ne regardent que le mode d'émission se reflètent dans le son lui-même en des différences que l'oreille peut détecter et la spectrographie étudier. Nous ne dirons rien ici de ce difficile problème sur lequel l'examen des analyses suivantes nous obligera à revenir, et terminons cette



N.B. sur chaque graphique, on a figuré l'origine ainsi que l'unité choisie sur les axes (sauf en (VII ; 3,4) où l'on a figuré la 1/2 unité) ; les axes 1,3 sont horizontaux, les axes 1,2 sont verticaux (sauf en (XII ; 1,3) où l'axe 3 est vertical).

Figure 6.2

description partielle du système des traits pertinents par un tableau emprunté à M. & N. Dans ce tableau 0 signifie NON et 1 signifie OUI ; exceptée la colonne place, où on note : 0 labiale, 1 dentale, 2 gutturale.

Tableau 6-1

Les traits pertinents des consonnes selon M. & N.

| Consonne | Sonore | Nasale | Fricative | Longue | Place |
|----------|--------|--------|-----------|--------|-------|
| p | 0 | 0 | 0 | 0 | 0 |
| t | 0 | 0 | 0 | 0 | 1 |
| k | 0 | 0 | 0 | 0 | 2 |
| f | 0 | 0 | 1 | 0 | 0 |
| θ | 0 | 0 | 1 | 0 | 1 |
| s | 0 | 0 | 1 | 1 | 1 |
| ∫ | 0 | 0 | 1 | 1 | 2 |
| b | 1 | 0 | 0 | 0 | 0 |
| d | 1 | 0 | 0 | 0 | 1 |
| g | 1 | 0 | 0 | 0 | 2 |
| v | 1 | 0 | 1 | 0 | 0 |
| ϕ | 1 | 0 | 1 | 0 | 1 |
| Z | 1 | 0 | 1 | 1 | 1 |
| 3 | 1 | 0 | 1 | 1 | 2 |
| m | 1 | 1 | 0 | 0 | 0 |
| n | 1 | 1 | 0 | 0 | 1 |

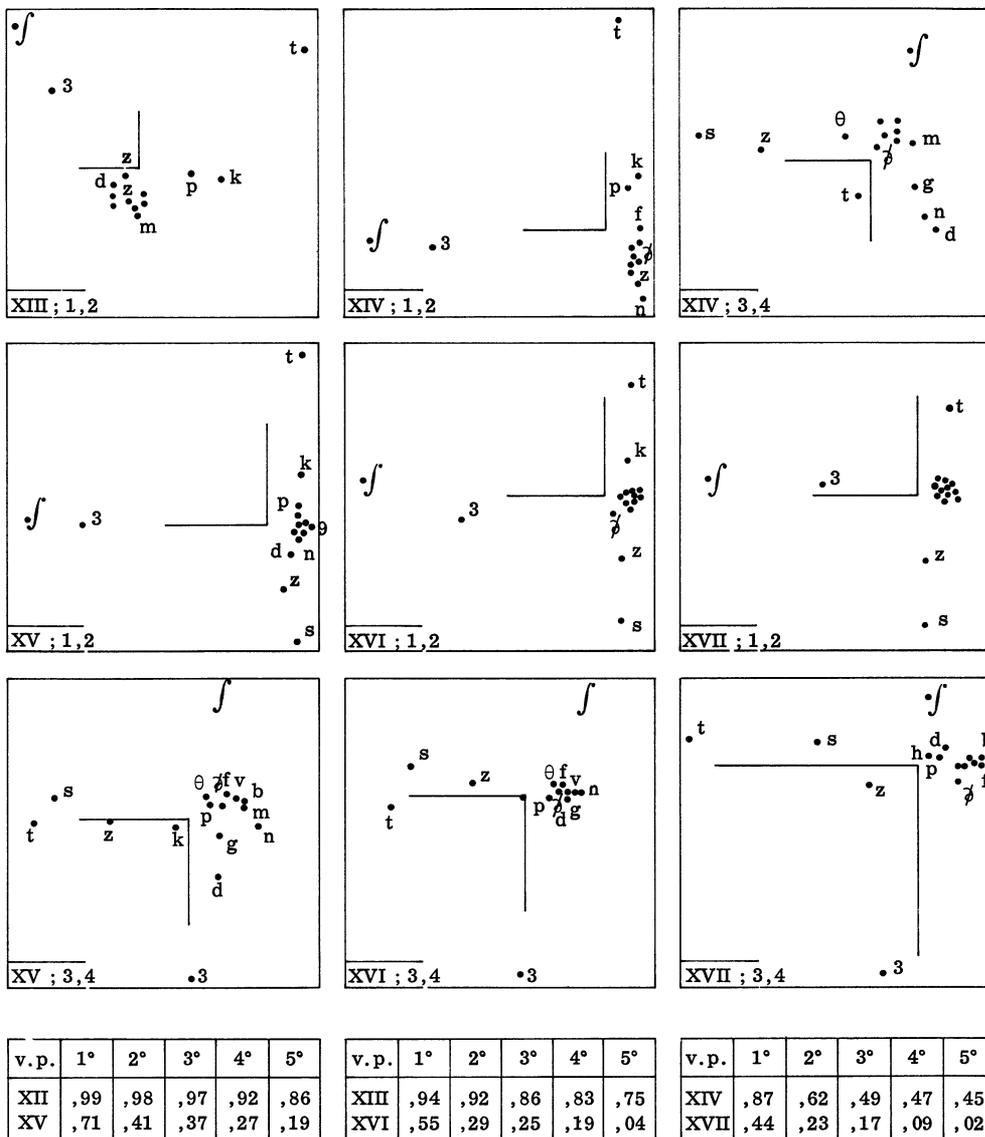
6.4 - A partir du tableau VII, considérons maintenant la série VIII, IX, X, XI, XII. La bande passante du filtre, toujours limitée inférieurement à 200 cps, voit sa limite supérieure passer de 300 cps (pour VII) à 5000 cps (pour XII). Les conditions s'améliorent jusqu'à s'approcher de l'optimum (200 - 6500 cps, 12 dB, condition VI) atteint dans les expériences de M. & N. : aussi dans les tableaux XI et XII les confusions sont-elles fort rares. Examinons les résultats dans le plan des deux premiers facteurs. L'opposition majeure est entre les nasales {m,n} et le reste des consonnes ; celles-ci étant scindées en trois groupes : les sourdes occlusives {p, t, k} ; les sourdes fricatives {f, θ, s, ∫} ; et les 7 sonores orales. Dans le cas VII, IX, X, XI et XII les nasales sont séparées par le premier facteur. Dans les cas VIII, IX, X et XI, le deuxième facteur sert à distinguer les trois groupes de consonnes orales (i.e. non nasales). Mais, dans le cas XII, le ∫ qui n'est l'objet que de quatre confusions en tout comme stimulus et comme réponse, se trouve sur le deuxième axe opposé seul aux 15 autres phonèmes qui sont groupés tout près de l'origine : on reconnaît là un exemple de facteur associé à une partition, (cf. e.g. supra n° 4 ; c'est d'ailleurs aussi à une partition - en trois classes - que sont associés les facteurs 1 et 2 issus du tableau VIII à XI). Cependant le troisième facteur issu de l'analyse XII diffère peu du 2ème facteur issu des autres analyses de la sé-

rie (à ceci près que le groupe des fricatives sourdes est contigu au groupe des sonores orales) ; aussi a-t-on représenté dans le plan 1-3 les résultats de l'analyse XII.

Comment expliquer la nouvelle organisation tripartite des consonnes orales, un peu divergente des conceptions linguistiques reçues ? On pourrait conjecturer, e.g., que dans la bande 200 - 1000 cps, la puissance totale des fricatives sourdes {f, θ, s}, est intermédiaire entre celle des occlusives sourdes (p, t, k et celle des sonores ; tandis que dans la bande plus étroite 200 - 300 (condition VII) l'opposition est entre les sourdes (occlusives ou non) qui n'ont presque rien dans cette bande de basse, et les sonores qui y ont une puissance notable (sons grasseyants, avec bourdonnement des cordes vocales ; cf. supra, perception des sonores). Cependant selon Strevens (1960) le spectre du bruit des fricatives ne débute qu'au delà de 1000 cps : il faut donc admettre que la reconnaissance des fricatives sourdes repose ici sur une autre propriété, telle que le déplacement d'un format de la voyelle...

Passons aux facteurs 3 et 4 ; quand la bande passante s'élargit, le taux de confusion diminue, ces facteurs deviennent de moins en moins sûrs. Considérons le cas VIII : la partition des sonores orales en occlusives et fricatives a disparu ; elle reparait toutefois dans le cas IX, où le plan (3, 4) révèle les mêmes îlots que dans le cas VII. Nous verrons souvent dans la suite que la forte affinité du v avec le b brouille l'opposition occlusive-fricative à l'intérieur des sonores : déjà dans l'analyse VII sur le plan (3, 4), le v est déplacé vers {b, d, g}.

6.5 - Examinons la série XII, XIII, XIV, XV, XVI, XVII, des filtres passe-haut. Pour Miller & Nicely (p. 350 loc. cit.) : the errors do not cluster or fall into obvious patterns in the confusion matrix, but seem to distribute almost randomly. Les auteurs qui ont dépouillé les tableaux en se fiant aux oppositions reçues des linguistes (e.g. disons en calculant combien il y a de paires occlusives-fricatives, ou longues brèves ayant fait l'objet de confusions) remarquent toutefois que : duration holds up somewhat better because |s|, |ʃ|, |Z| and |ʒ| (les longues pour M. & N.) are characterized in part by considerable high frequency energy. Ici, l'analyse factorielle, est à même d'extraire du hasard apparent des confusions une structure dont la stabilité nous assure qu'elle est significative. Les quatre premiers facteurs sont, à des variations mineures près, communs aux trois tableaux XVII, XVI, XV. En bref disons que toutes les consonnes sont confondues à l'origine, exceptées les quatre longues (de M. & N.) {s, Z, ʃ, ʒ} et aussi la dentale occlusive t qui selon M. & N. (page 347) est caractérisée par un bruit de haute fréquence. On peut dire que, sur le premier facteur, {ʃ, ʒ} gutturales s'opposent à {s, Z, t} (dentales) : toutefois, il n'est pas permis pour autant d'interpréter ce facteur comme donnant le point d'articulation des sons perçus. Il est incontestable que le point d'articulation est un trait pertinent essentiel de l'acte d'émission, mais il n'en résulte pas nécessairement qu'à l'audition les sons apparaissent rangés en trois classes suivant le point d'articulation donné par le tableau 6 - 1. Au contraire, pour M. & N. (p. 347), particulièrement dans les conditions d'écoute per-



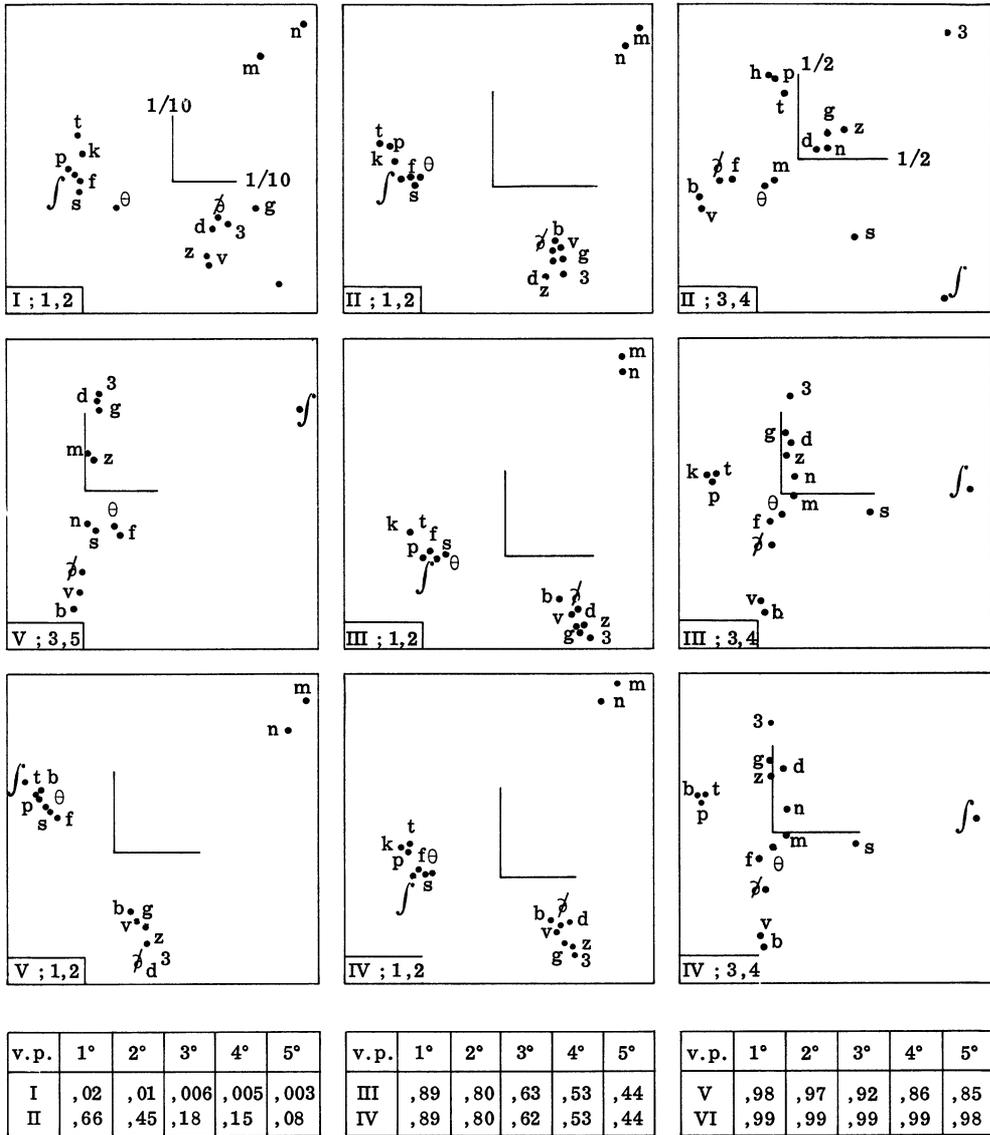
N.B. sur chaque graphique, on a précisé l'origine, ainsi que l'unité choisie sur les axes ; les axes 1,3 sont horizontaux, les axes 1,2 sont verticaux.

Figure 6.3

turbées qui sont celles de leurs expériences, la reconnaissance du point d'articulation serait faite différemment au sein de chaque classe de consonnes (e.g. les occlusives sourdes, ou les occlusives sonores orales). Aussi ne faut-il pas demander que le point d'articulation apparaisse, comme un facteur. (cf. toutefois, infra n° 6-5, facteur IV⁴). Dans les cas XIII et XIV la bande passante s'élargit, le taux de confusion décroît ; sur le plan des deux premiers facteurs on ne distingue plus que $\{s, \int\}$ et t, opposés à toutes les autres consonnes ; toutefois, pour l'analyse XIV le groupe $\{s, z\}$ est encore visible dans le plan (3ème, 4ème). Au reste, il faut souligner que dans les cas XII et XIV la bande passante étant également altérée à ses deux extrémités des causes d'erreurs fort diverses se conjuguent, rendant plus difficile l'analyse des confusions.

6.6 - Reste la série, I, II, III, IV, V, VI, où le son, non-filtré (si non par les écouteurs de bande passante 200 - 6500) est mêlé à un bruit blanc d'intensité variable. A l'extrémité VI, les conditions sont optima ; il y a si peu d'erreurs que l'analyse factorielle globale n'est pas sûre : tout ce qu'on peut noter sur le tableau VI c'est que des confusions systématiques, subsistent à l'intérieur des îlots $\{p, k\}$, $\{f, \theta\}$, $\{d, g\}$, $\{b, v, \beta\}$. Ceci dit, considérons les analyses I à V. Dans le plan 1-2, les résultats sont les mêmes que pour l'analyse VII : le premier facteur est la sonorité : il sépare les sourdes des sonores (orales ou nasales) : le 2ème facteur peut être appelé la nasalité : quasi-nul pour les sourdes il sépare les sonores nasales $\{m, n\}$ des autres sonores $\{b, d, g, v, \beta, Z, \int\}$ (à moins que, sans chercher d'interprétation individuelle de facteurs ; on se borne à dire que les facteurs 1 et 2, ensembles, séparent les 16 consonnes en trois classes : sourdes, sonores orales et sonores nasales). Toutefois dans le cas I, où les erreurs sont très nombreuses (et où nous l'avons dit plus haut au n° 6.2, on est au seuil de la signification) l'îlot des sourdes et celui des sonores sont très dispersées ; tandis que dans le cas V où au contraire les confusions sont très rares donc les facteurs peu stables (on tend vers le cas d'une matrice diagonale pour laquelle seule existe la valeur propre multiple 1, toute fonction pouvant être prise pour facteur) il y a une légère rotation de l'ensemble des îlots dans le plan 1,2.

Dans le plan 3,4, l'analyse I n'offre rien de significatif ni d'interprétable. Les analyses III et IV fournissent des résultats presque identiques. Les sonores nasales sont à l'origine. Les sourdes s'échelonnent approximativement le long du 3ème axe : d'abord les occlusives $\{p, t, k\}$, qui coïncident presque ; puis les fricatives dans l'ordre $\{f, \theta, s, \int\}$. Les sonores orales sont de même approximativement rangées sur le quatrième axe. L'ordre en est : $\{g, d, Z, \beta, v, b\}$; occlusives et fricatives sont mêlées, mais il est satisfaisant de noter que les sonores sont rangées d'après leur point d'articulation d'arrière en avant : les gutturales $\{g, \int\}$, puis dentales $\{d, \beta, \beta\}$, puis labiales $\{v, b\}$. L'analyse II offre dans le plan (3,4) une configuration peu différente de celle des analyses III et IV ; mais il y a une rotation de l'ensemble d'environ 60°. Quant à l'analyse V le quatrième facteur, quasi-nul pour toutes les consonnes orales, ne fait que séparer le m du n : en effet m et n, d'abord distingués dans le plan (1, 2) parce qu'il ne sont confondus que 2 fois avec d'autres sons, n'ont entre eux-mêmes que cinq confusions ce qui explique (effet de partition) qu'un facteur leur soit attribué.



N.B. sur chaque graphique, on a figuré l'origine, ainsi que l'unité choisie sur les axes (sauf en (I ; 1,2) où l'on a figuré le 1/10° d'unité et en (II ; 3,4) où l'on a figuré la 1/2 unité : les axes 1,3 sont horizontaux, les axes 2,4 (ou 5) sont verticaux.

Figure 6.4

Mais dans le plan (3, 5) on reconnaît la configuration d'ensemble qui pour les analyses III, IV (et II) apparaît sur le plan (3, 4).

Que chez les sonores l'opposition occlusive fricative n'ait en présence de bruit qu'un rôle mineur, peut s'expliquer par ce que les fricatives sonores ne se distinguent des occlusives sonores qu'en ajoutant aux vibrations musicales (sonorité) un chuchotement qui se perd dans le bruit blanc. Les occlusives sourdes continueraient de former une classe distincte, soit de par leur brièveté, soit parce que, en l'absence de sonorité le chuchotement fricatif se distingue mieux du bruit.

Nous espérons que le lecteur statisticien aura suivi avec attention notre analyse des belles données de Miller & Nicely. En effet, d'une part l'acoustique de la parole est intéressant en elle-même. D'autre part les hypothèses suggérées par les 17 analyses sont un exemple de ce que l'examen mathématique des données peut apporter aux sciences expérimentales. Les statistiques calculées par M. & N. en restant dans le cadre structural des oppositions définies par les linguistes, laissaient échapper des phénomènes qu'ont révélé notre analyse des tableaux, menée sans hypothèse à priori. Quant à S. Johnson qui utilise des méthodes de classification mathématique, il n'a publié jusqu'ici que l'analyse du tableau VII, le mieux conforme, on l'a vu, à la structure reçue des phonologues ; mais il affirme (1967) que les 16 autres analyses lui ont donné des résultats surprenants encore que cohérents entre eux, et annonce que l'explication détaillée en sera donnée par R. N. Shepard dans un rapport que nous n'avons pas reçu. La cohérence interne des classifications ne nous étonne nullement car elle est conforme à ce que nous avons vu plus haut : toutefois sans minimiser la valeur des programmes de classification que nous utilisons aussi, nous pensons que certaines des informations précises mises en évidence par l'analyse factorielle ne s'exprimeraient qu'imparfaitement dans une hiérarchie taxinomique. Enfin, signalons que K. V. Wilson (1961) a publié une analyse factorielle du tableau II. Cet auteur utilise les nombres de confusions $k(i, i')$ et $k(i', i)$ pour calculer une distance entre i et i' ; puis suivant la méthode générale d'analyse d'un ensemble muni de distances (méthode appelés aux U.S.A. Multidimensional Scaling) il donne une représentation euclidienne du système des 16 phonèmes. Wilson utilise deux formules de distance. L'une de ces formules (due à Wilson lui-même) :

$$1 - d(i, i')^2 = k(i, i') [(1/k(i)) + (1/k(i'))] / [k(i, i)/k(i) + (k(i', i)/k(i'))],$$

conduit à des résultats assez semblables à ceux de l'analyse des correspondances. Dans le plan des deux premiers facteurs, on a trois îlots : sourdes, nasales, et sonores orales. Sur les axes trois et quatre on a les séries :

$$\{ \int_3 Z g d \int n t p k s m \dots \theta f \int v b \}$$

$$\{ p t k \dots \int b m v n g d Z_3 \dots f \theta \int s \}$$

(où les points de suspension marquent une certaine distance).

7 - COMPARAISON ENTRE PAIRES DE SIGNAUX DE L'ALPHABET MORSE

7.0 - L'expérience de E.Z. Rothkopf (1957) que nous abordons maintenant a déjà par la richesse de ses résultats intéressé les spécialistes de l'analyse des données (cf. e.g. Shepard, 1963). En vue de mesurer les similitudes entre stimulus, E.Z. Rothkopf proposa à 598 sujets des suites de signaux Morse présentés par paires : à chaque paire les sujets (qui avaient été choisis pour n'avoir aucune connaissance préalable du code télégraphique et ne recevaient aucune indication sur la valeur alphanumérique des signaux) devaient simplement noter sur une feuille de réponse si les deux signaux leur paraissaient identiques ou différents. Les sujets étaient divisés en huit groupes de taille inégale : chaque groupe était soumis à une dictée différente ; mais chaque dictée comprenait, rangées au hasard 351 (i.e. $36 \times 35/4$) paires de signaux distincts et les 36 paires formées de deux signaux (lettres ou chiffres) identiques. En définitive, du dévouement des feuilles de réponse Rothkopf a déduit un tableau 36×36 donnant pour chaque paire (i, i') de signaux (lettres ou chiffres) la fréquence $k(i, i')$ avec laquelle les sujets jugeaient identiques les deux signaux présentés dans l'ordre (i, i') : de façon précise, si la paire (i, i') a été présentée à $n(i, i')$ sujets parmi lesquels $s(i, i')$ ont estimé les deux signaux identiques, le tableau de Rothkopf contient les valeurs approchées à une unité près des nombres $100 s(i, i')/n(i, i')$. Si l'expérience avait été conduite avec une symétrie rigoureuse, les $n(i, i')$ (du moins pour $i \neq i'$ seraient tous égaux entre eux ; et la matrice $k(i, i')$ serait, à un coefficient constant près, une véritable matrice de confusion. Il n'en est pas exactement ainsi : mais, chaque paire ayant été présentée à 150 sujets environ (aux 598 sujets s'il s'agit d'une paire (i, i) de deux signaux identiques), on peut approximativement accorder aux nombres entiers du tableau de Rothkopf la même confiance qu'à des nombres de confusions (et même une confiance plus grande, puisque e.g. 4 % correspond à 6/150 (Il faut toutefois signaler une anomalie fâcheuse dont Rothkopf lui-même fait état : au cours de la dictée, le taux de confusion décroît, car les sujets s'accoutument à leur tâche de comparaison. Or chaque paire de signaux ne figure que dans deux dictées : le taux de confusion trouvé dépend donc quelque peu du rang de ces deux paires : à similitude égale il est d'autant plus faible que la paire a été présentée plus tard dans les dictées). Nous avons donc soumis le tableau de Rothkopf à une analyse de correspondance. Les cinq premières valeurs propres sont : 0,256 ; 0,196 ; 0,153 ; 0,093 ; 0,063. Nous examinerons ici successivement les trois premiers facteurs extraits.

7.1 - Pour interpréter le premier facteur, rappelons comment sont émis les signaux Morse : On sait qu'ils se composent de deux sortes d'éléments : les points et les traits. Dans l'expérience de Rothkopf le rythme est rapide : un point dure 5/100 de seconde ; un trait 15/100 ; l'intervalle entre deux éléments d'un même signal est de 5/100 ; les deux signaux d'une même paire sont séparés par un silence de 1,4 sc ; le silence entre deux paires successives dure 3 sc. Les sujets n'ont donc pas le temps de noter la composition des signaux, ils doivent juger des similitudes d'après une impression globale.

Sur la figure 7.1 en a représenté le premier facteur (à l'échelle de 10 cm pour 1) en traçant, en face de chaque signe la suite de traits et de points dont il se compose. (La matrice de Rothkopf n'étant pas exactement symétrique, chaque signe devrait apparaître deux fois : comme 1er élément

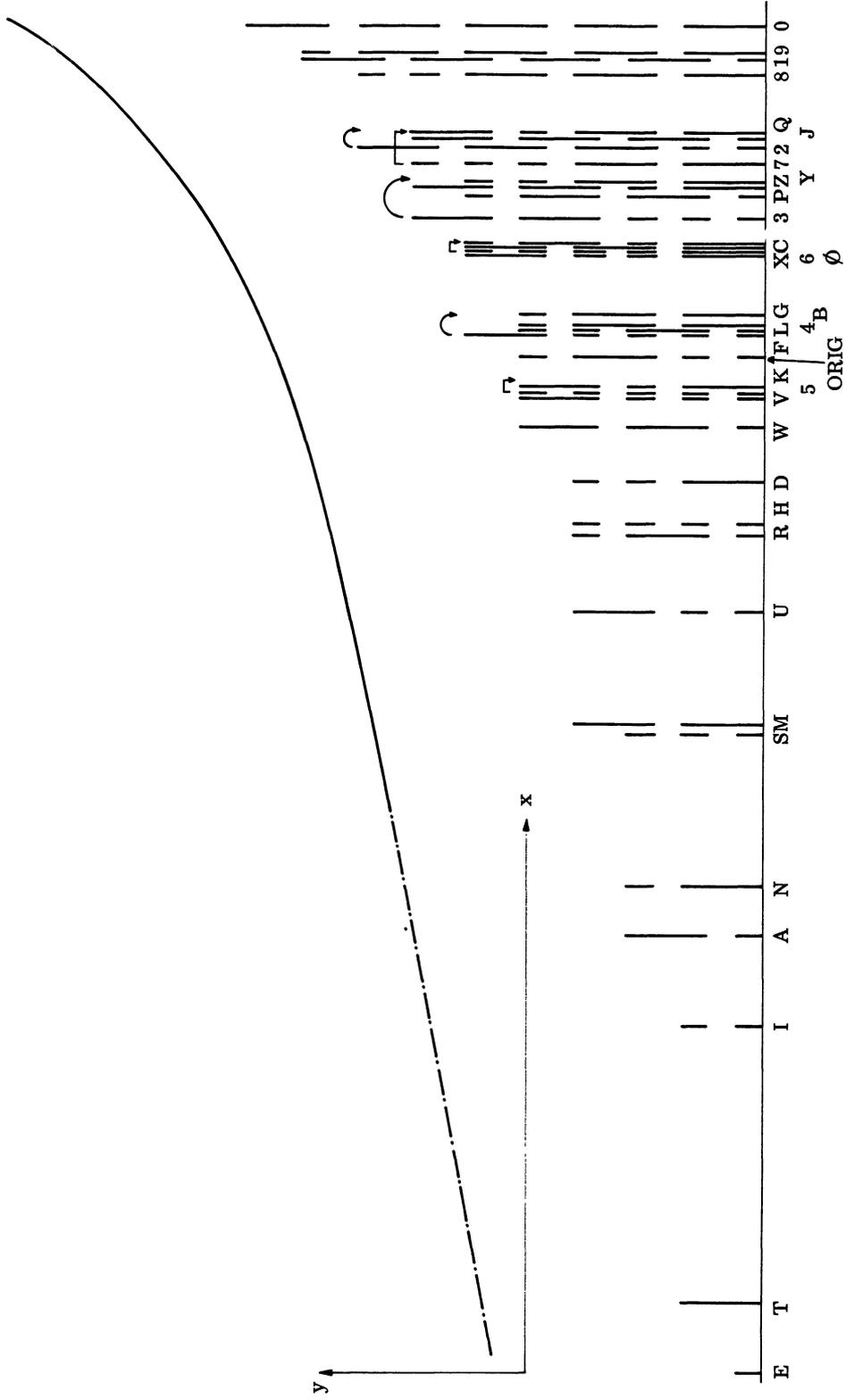


Figure 7.1 - Analyse de l'expérience de Rothkopf : facteur et courbe de durée subjective

de la paire, et comme deuxième élément ; dans cette figure-ci, on a donné à chaque signe l'abscisse moyenne de ses deux représentations, (sur les figures 7.2 et 7.3 on a noté à la fois deux points en les joignant d'un trait). Il apparaît qu'à quatre exceptions près (les exceptions concernent les signes 4, 3, Y et 2 : le signe 4 est décalé de 3 rangs, se trouvant à gauche de L, B et G ; le signe 3 est à gauche de P et Z ; Y est à gauche de Z ; 2 est à gauche de J et Q ; parmi ces inversions, celle de 4 et L et celle de Y et 2, ne sont sûrement pas significative (cf. infra) car elles portent sur 0,3 mm et 0,1 respectivement et ont dû être exagérées pour apparaître sur la figure 7.1 ; sur cette figure, les exceptions sont marquées d'un crochet en demi-cercle qui renvoie le signal à la position que lui assignerait sa durée), l'ordre des 36 signes sur le premier axe coïncide avec celui de leurs durées totales. Un tel résultat ne peut évidemment être fortuit. Si d'ailleurs on tente de chiffrer la précision avec laquelle sont calculés les facteurs, on peut dire que selon les aléas d'expérience un point i est placé sur l'axe factoriel avec un écart-type de l'ordre de $k(i)^{-1/2}$. Comme une ligne ou colonne du tableau de Rothkopf a un total d'environ 600 (nombre que l'on peut, cf. ci-dessus, multiplier encore par 1,5 pour avoir le nombre d'événements par ligne ou colonne), l'erreur sur un point doit être de l'ordre de quelques centièmes, c'est-à-dire sur la figure de quelques millimètres : de semblables erreurs ne peuvent pas détruire d'ordre des points.

7.2 - L'interprétation psychologique est pour nous la suivante : les sujets, non familiers avec les signaux sont dépassés par la tâche de les comparer ; parmi toutes les ressemblances qu'ils peuvent entrevoir, l'égalité de durée prédomine : de même que l'homonculus (cf. n° 3) est lié à une échelle de distance sensorielle, le premier facteur est une échelle de durée subjective des signaux. Considérons les lignes (suites de traits et de points séparés par des intervalles) qui sur la figure représentent les 36 signaux : ces lignes débutent sur l'axe rectiligne, et leurs extrémités se placent sur une courbe qui a approximativement l'allure d'une exponentielle. Pour être précis, appelons axe des abscisses x l'axe factoriel ; et axe des ordonnées y l'axe perpendiculaire au premier. L'ordonnée y de l'extrémité de la ligne figurant un signal est la durée réelle ; tandis que les variations d'abscisses correspondent à des variations de durée subjective. La relation approximative $x \approx \log y$, nous rappelle la loi dite de Weber Fechner selon laquelle la sensation (ici le facteur x) est fonction linéaire du logarithme de l'excitation (ici la durée réelle) : cette loi s'introduit naturellement si l'on mesure la variation de la sensation en nombre de seuils, et que l'on admet que le seuil de discriminabilité entre stimuli est obtenu pour une variation relative, $\Delta y/y$, du stimulus y - ici une durée - indépendante de y . Dans cette optique théorique, il n'y a pas d'origine qui s'impose sur l'axe des x (quand y varie de 0 à l^∞ , $x = \log y$ varie de $-\infty$ à $+\infty$) : c'est pourquoi nous prenons soin de parler de variation de durée subjective plutôt que de durée objective. En fait il y a bien un zéro absolu de durée subjective, correspondant à la durée minima produisant un effet ponctuel mais ce phénomène est en dehors de la partie régulière logarithmique de l'échelle. Dans le cas présent, le zéro et l'unité de l'axe factoriel sont fixés par la condition de normalisation que le facteur ait moyenne nulle et variance unité sur l'ensemble des stimuli (signaux Morse) étudiés. La relation logarithmique, assez bien vérifiée pour les durées réelles grandes (à partir de durée égales à

points), n'a plus de sens pour les signaux brefs (E, T, I). Pour tracer la courbe représentative $y = \varphi(x)$, de la durée réelle comme fonction φ du facteur x de durée subjective) on a utilisé dix points obtenus en prenant les moyennes des extrémités des lignes sur les 10 groupes de signaux de durée 1(A), 3(T, I), 5(A, N, S)... jusqu'à 19 (0, signal formé de 5 traits dont la durée totale, intervalles compris est celle de 19 points). Sur la figure 7.1, la courbe $y = \varphi(x)$ est tracée décalée dans la direction de l'axe des ordonnées, afin qu'elle ne se mêle pas aux lignes de traits et de points représentant les signaux. Pour la même raison, la direction de l'axe des y est indiquée par une flèche ne passant pas par l'origine, mais située à gauche de la feuille ; l'origine vraie de l'axe factoriel est indiquée par une courte flèche perpendiculaire à cet axe et marquée orig.

[Signalons que dans une précédente publication (1969) nous avons commis l'erreur d'attribuer à l'intervalle entre deux éléments d'un signal la durée du trait au lieu de celle du point. L'interprétation du premier facteur était cependant semblable à celle qui résulte d'un calcul exact des durées à deux différences près. D'une part l'ordre des signaux sur le 1er axe présentait un bien plus grand nombre d'inversions. Mais d'autre part la durée calculée : $4p + 6t - 3$ (où p est le nombre de points et t le nombre de traits d'un signal), au lieu de la vraie valeur $2p + 4t - 3$, se corrigeait à .94 avec le 1er facteur. En réalité, la durée est une fonction non linéaire du premier facteur, ce qui est, somme toute, plus conforme aux lois psychophysiques requises].

7.3 - L'examen attentif des groupes de signaux dont la durée réelle est la même suggère une règle qui sur la figure est suivie sans exception Soient deux signaux l'un composé de p points suivis de t traits l'autre de t traits suivis de p points : c'est ce dernier qui a la durée subjective (plus exactement l'abscisse sur le 1er axe) la plus grande. On le vérifiera pour les paires (A, N) ; (U, D) ; (W, G) ; (V, B) ; (4, 6) (3, 7) ; (2, 8) ; (1, 9). Il n'est évidemment nullement démontré qu'il s'agisse d'une loi psychophysique de la perception de la durée. Mais l'analyse factorielle nous semble du moins suggérer des hypothèses, donc des expériences : il conviendrait d'étudier la durée subjective de signaux formés d'une suite d'éléments sonores soit de plus en plus brefs, soit de plus en plus longs : l'accélération ou le ralentissement du rythme influe-t-il sur l'estimation de la durée ? Peut-être ; mais une explication plus simple se présente à l'esprit : le dernier élément entendu paraît toujours plus long, même s'il est bref. En effet considérons les 8 groupes formés de signaux de même durée : (T, I), (A, N, S), (M U R H D), (W V 5 K F L B G), (4 X 6 Ø C P Z), (3 Y 7 3 Q), (2, 8), (1, 9) : il est de règle que, sur la figure 7.1, les signaux de chaque groupe terminés par un point soient à droite de ceux du même groupe terminés par un trait : cette règle ne souffre que trois exceptions (marquées sur la figure 7.1 par un crochet carré) concernant les signes 5, 6 et 7 ; et seule l'exception relative au signe 7 est significative. De façon précise puisque par exemple, les trois signaux (R H D) qui durent comme 7 points et se terminent par un point, se placent entre le groupe (M U), (durée 7, terminés par un trait), et le groupe (N V K), (durée 9, terminés par un trait), on peut dire qu'à durée réelle égale un signal terminé par un point paraît durer un point de plus qu'un signal terminé par un trait. Il est curieux de noter ici que les anciens, afin de justifier la règle de versification latine qui permet de placer à la fin d'un vers indifféremment une syllabe brève ou

longue (alors que la mesure des vers est définie par l'ordre des quantités des syllabes) disaient qu'avant un silence, toute syllabe paraît longue. Saint Augustin (de Musica L. IV C.2) s'exprime ainsi : "...et ob hoc spatium temporis in ipsa quiete largissimum nihil distat quae ibi syllaba locetur; nonne illud est consequens, ut ipsa ultimae syllabae indifferentia, quae propter largum spatium conceditur, ad id proficiat, ut sive ibi brevis syllaba sive longa sit, eam sibi antes pro longa vindicent". (...et à cause du très grand espace qu'offre le silence, peu importe quelle syllabe s'y loge ; n'en résulte-t-il donc pas qu'indifféremment, que la dernière syllabe soit brève ou longue, l'oreille la prenne pour longue).

Quoiqu'il en soit de la règle selon laquelle l'ultime élément d'un signal paraît toujours long (plus précisément, nous avons proposé ci-dessus que le point ultime semble durer 2 points), on peut vérifier sur le tableau de Rothkopf (reproduit par Shepard 1963) que les confusions (exactement : les jugements d'identité) sont près de deux fois plus nombreuses sur les paires (x., x -) que sur les paires (.x, - x) (ou on a noté x une suite quelconque d'éléments). On a les chiffres suivants : à gauche, après chaque paire (x. x-) on a donné, séparés par le signe +, d'abord le taux de confusion correspondant à l'ordre de présentation (x. x-) puis celui pour l'ordre (x- x.); à droite, il s'agit des paires (.x -x) et à gauche des paires (-x .x).

| | |
|---|---------------------------------------|
| I..A.-64 + 46 ; N-.M--59 + 62 ; | I..N-.30 + 13 ; A.-M--25 + 24 ; |
| S...U.-56 + 51 ; R.-.W.--62 + 61 ; | S...D-.30 + 29 ; U.-.K--44 + 23 ; |
| D-.K--81 + 73 ; G--.Ø---82 + 76 ; | R.-.G-.26 + 21 ; W.--Ø---26 + 25 ; |
| H...V...-59 + 39 ; B...X...-84 + 64 ; | H...B...45 + 21 ; V...-X...-57 + 17 ; |
| G-.Y--82 + 62 ; P.--.J.---78 + 63 ; | F...C--.33 + 29 ; L...Z---.59 + 51 ; |
| Z--.Q--.72 + 63 ; | |
| 5.....4....-69 + 42 ; 9----.O-----78 + 81 ; 5.....6-....42 + 17 ; 1.----O-----55 + 50 ; | |

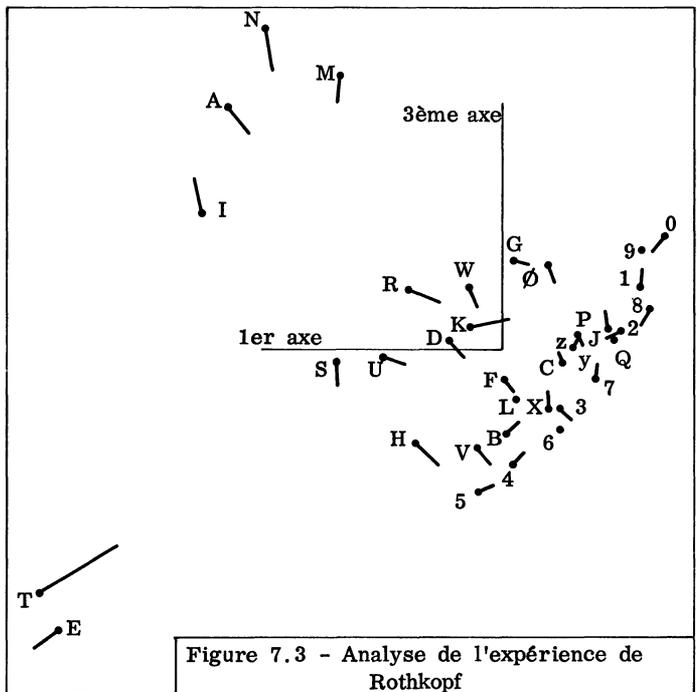
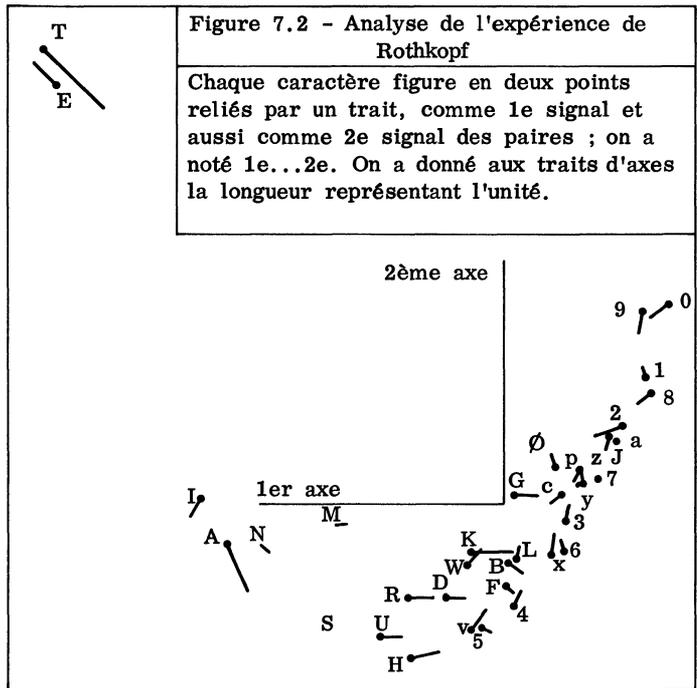
Plus précisément ce tableau fait apparaître les inégalités :

$$k(x., x-) > k(x-, x.) \gg k(x., -x) \gg k(-x., .x) :$$

Mais seule l'analyse factorielle (ou le raisonnement plausible de Saint Augustin) suggèrent que c'est (x.) qui paraît être (x-), plutôt que (x-) qui paraît être (x.).

7.4 - L'interprétation du 2ème facteur dépend de celle du premier : dans le plan des axes (1,2) en effet, (cf. fig. 7.2), les points se placent approximativement sur une parabole. On a donc ici un nouvel exemple de l'effet Guttman : le 2ème facteur est une fonction quadratique du premier. Ceci s'explique bien si l'on se rapporte à la situation limite de signaux dont on ne comparerait que les durées. La matrice de confusion ne contiendrait alors que des zéros, excepté dans une bande diagonale ; dans ce cas il y a bien effet Guttman. Par rapport au schéma parabolique, la figure du plan (1, 2) présente des déviations systématiques qu'on interprétera avec le 3ème facteur.

7.5 - Après la durée totale, c'est sans doute le nombre total d'éléments (traits ou points) qui est le caractère par lequel deux signaux successifs sembleront le mieux se reproduire l'un l'autre, à celui qui tente de les



comparer. Il est cependant mathématiquement impossible qu'après la durée totale, le nombre d'éléments sorte aussi comme facteur. En effet, deux facteurs doivent être non-corrélés : soit $F_1(s)$ et $F_2(s)$ deux facteurs, supposés réduits à avoir moyenne nulle, on a (en notant S l'ensemble des stimuli, et $k(s)$ le poids relatif de chacun d'eux) :

$$\sum \{F_1(s) F_2(s) k(s) \mid s \in S\} = 0$$

Or au sein de l'alphabet Morse la durée totale d (exprimée par exemple en durée de points par la formule considérée plus haut : $2p + 4t - 3$) sans être exactement une fonction linéaire du nombre total d'éléments ($n = p + t$) n'en varie pas moins approximativement dans le même sens que lui : une fois réduits tous deux à avoir moyenne nulle d et n seront donc le plus souvent du même signe (tous deux sont simultanément soit grands, soit petits : e.g. quand n atteint sa valeur maxima 5, d est supérieur ou égal à 9, durée du signal chiffre 5 ; or 9 est à peu près la valeur moyenne de d). On voit donc, sans plus de calcul que n et d sont corrélés positivement ; et, de même $n(s)$ (nombre d'éléments) est corrélé au premier facteur $F_1(s)$ de durée subjective, que nous avons extrait. Ceci exclut que le nombre total d'éléments étant corrélé au 1er facteur, puisse être le 3ème facteur ; bien plus le troisième facteur devra aussi être orthogonal (non corrélé) au second qui étant fonction parabolique de la durée est lui aussi lié à $n(s)$. Le nombre total $n(s)$, bien qu'ayant après la durée un rôle prépondérant, ne pourra pas apparaître sur les graphiques en aussi belle place qu'on le souhaiterait.

Soient les 5 groupes de signaux (à 1, 2, 3, 4, 5 éléments respectivement), (E, T), (I, A, N, M) etc., chacun rangé dans l'ordre de la durée croissante (en tenant compte des variations, régulières ou fortuites, de la durée subjective, constatées en examinant le 1er facteur). Quelle pourrait-être la représentation idéale de l'ensemble (ainsi divisé en 5 séries ordonnées) des 36 signaux ? Dans le plan rapporté à deux axes orthogonaux ox oy (issus du centre de gravité du nuage des 36 points), chaque signal s aurait des coordonnées $x(s)$ et $y(s)$ telles que :

1/ $x(s) = F_1(s)$, facteur de durée subjective déjà reconnu

2/ $\sum \{x(s) y(s) k(s) \mid s \in S\} = 0$: les coordonnées orthogonales x et y sont non corrélées sur l'ensemble des signaux ; ox et oy sont donc (d'après une propriété géométrique générale) les axes principaux d'inertie du nuage des 36 signaux.

3/ $n(s) = a x(s) + by(s) + c$: autrement dit $y(s)$ est la combinaison linéaire de $x(s)$ (durée subjective) et de $n(s)$ (nombre d'éléments) qui est non corrélée à $x(s)$ (combinaison unique à un coefficient près) ; ou encore : durée subjective et nombre d'éléments forment, sinon un système de coordonnées orthogonales, du moins un système de coordonnées obliques ; les cinq classes de signaux (définies par la valeur de $n(s)$), se rangent en chapelets sur les droites.

4/ il resterait à s'assurer que $y(s)$ - candidat selon nous à être le troisième facteur - n'est pas corrélé à $F_2(s)$; mais nous n'examinerons pas ici ce point particulier.

Sur la figure 7.4, on a tenté de satisfaire approximativement aux conditions 1/ ; 2: et 3/. A cette représentation hypothétique, comparons les

E.T _
 I.. A..N..M _ _
 S...U...R...D...W...k...G...Ø... _ _ _
 H...V...B...F...L...X...C...P...Z...Y...J...Q... _ _ _ _
 5....4.....6.....3.....7.....2.....8.....1.....9.....0..... _ _ _ _ _

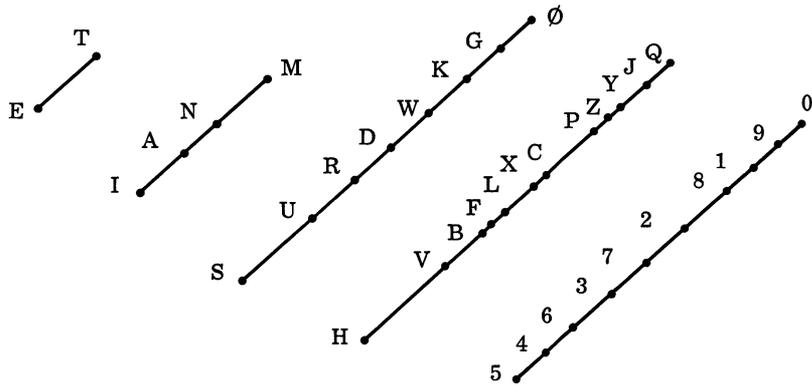


Figure 7-4 : Représentation systématique des séries de caractères contenant le même nombre d'éléments (cf. Figure 7-3 et 7-5)

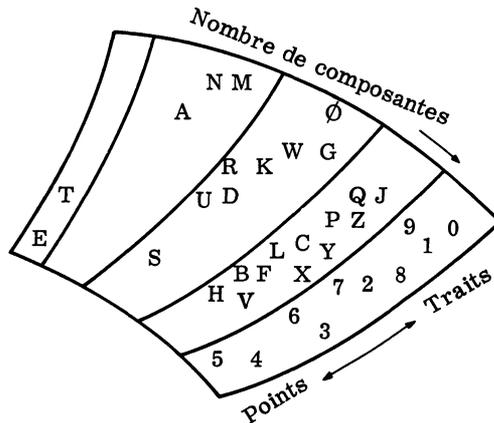


Figure 7-5 : Résultats de l'analyse des proximités ; d'après Shepard (1963)

représentations des plans (1, 2) et (1, 3) (cf. fig. 7.2, 7.3). Les groupes (E, T) et (I, A, N, M) sont bien détachés ; et celui-ci est dans l'ordre attendu. Sur le plan (1, 2), les trois autres groupes sont soudés, comme appliqués à la parabole de régression. Sur le plan (1, 3), le groupe (S, U, R, D, W, K, G, Ø) se sépare ; mais on ne peut distinguer la strate des lettres à quatre éléments de celle des chiffres (signaux à cinq éléments). Puisque le 2ème facteur est approximativement une fonction du premier (seul le M s'écarte notablement de la parabole de régression ou ligne moyenne du nuage) il est naturel de chercher dans le plan (1, 3) notre représentation globale du nuage des 36 signes. C'est donc au graphique du plan (1, 3) que nous comparons maintenant la représentation bidimensionnelle elle aussi, que R.N. Shepard (1963) a construite par l'analyse des proximités.

7.6 - Rappelons que Shepard considère comme un indice de proximité le pourcentage de confusion donné par le tableau de Rothkopf. Par exemple la moyenne de pourcentage de confusion en les signaux B et D présentés dans l'ordre BD et DB est environ 46 ; la moyenne des taux de confusions pour AG et CA est 5 : Shepard en conclut que la proximité subjective entre A et C est moindre que celle entre B et D. L'objet de l'analyse des proximités est de faire de l'ensemble des 36 signaux une carte (de dimension aussi faible que possible ; ici, plane) de telle sorte que la plupart des inégalités spécifiées par le tableau de Rothkopf soient satisfaites par les distances mesurées sur cette carte. Et il se trouve (remarquable découverte de Shepard, 1962) que ces conditions définissent suffisamment la carte cherchée ; et un algorithme itératif permet de construire celle-ci.

Sur la carte des 36 signaux publiée par R.N. Shepard (cf. fig. 7.5), l'organisation de notre figure 7.4 apparaît assez bien : les cinq séries sont rangées approximativement dans l'ordre que nous avons trouvé : la série des chiffres et celle des lettres à quatre éléments sont sur la figure 7.4, mieux étalées, mieux séparées que sur la figure 7.3 (plan 1.3). Les signaux (E, T-) sont moins à l'écart, ce qui serait plutôt un avantage ; mais ils semblent rejoindre la série (I..ANM) ; ce qui n'est ni satisfaisant du point de vue systématique (il convient de distinguer les signaux à un élément de ceux à deux éléments), ni conforme aux données de Rothkopf : la proximité entre E et T (56) est bien plus grande que celle entre T et I (15).

Pour R.N. Shepard, il n'y a pas sur la carte d'axes privilégiés : l'analyse des proximités ne peut donc, comme le fait l'analyse des correspondances, placer un premier axe sur lequel les signaux se projettent à 4 exceptions près dans l'ordre de leurs durées totales. Dans son interprétation, Shepard ne peut donc signaler le rôle prédominant dévolu, nous semble-t-il, à la durée totale du signal. Au lieu de cela, il emploie "dimension de longueur, dans un sens vague, pour signifier le nombre total d'éléments (abus de langage qui se justifie parce que, comme nous l'avons dit plus haut, le nombre $n(s)$ et la durée $d(s)$ sont corrélés entre eux). Et, comme on l'a écrit sur la figure 7.5, il attribue l'ordre interne des séries (E, T), (I, A, N, M), (S, U, D, R, K, W, G, Ø), à une opposition point/trait. Les deux dimensions psychologiques fondamentales seraient donc n et, disons, t/p . On trouve ici un cas de corrélation analogue à celui étudié en détail plus haut : sur une série telle que (S, U, D, R, K, W, S, Ø), où n est constant (3 dans notre exemple), la durée totale vraie d croit de 5 à 11, tandis que, simultanément, le rapport t/p croit de 0 à l'∞. Le choix des

deux dimensions psychologiques fondamentales apparaît ainsi comme un problème aussi indéterminé que celui de deux axes de coordonnées dans un plan : il y a ici plus de deux axes qui soient interprétables dans le langage courant. Pourtant, l'analyse de correspondance nous semble imposer la dimension durée totale qui est responsable des deux premiers facteurs (valeurs propres 0,256 et 0,196 ; suivies de valeurs propres bien moindres 0,153 ; 0,093 ; 0,063...). Et si, après R.N. Shepard, nous cherchons dans le tableau de Rothkopf "some useful picture of the way these 36 signals are processed in man", les paramètres durée d , et nombre d'éléments n , nous apparaissent bien être ceux dont on conçoit le mieux qu'ils soient chez l'homme l'objet d'une élaboration perceptive indépendante. Enfin, rappelons que l'examen attentif du premier facteur (n° 7.3) a suggéré l'hypothèse que, placé en fin de signal, le point semble avoir une durée doublée. Cette hypothèse appelle une vérification expérimentale ; mais on a vérifié déjà sur les données de Rothkopf qu'en position finale le point se confond tout particulièrement avec le trait.

8 - L'APPRENTISSAGE DU CODE MORSE

8.1 - Tandis que dans l'expérience de Rothkopf les sujets comparent, par paires, des signaux Morse dont ils ignorent la valeur ; dans l'expérience de Plotkin (1943, p. 11) analysée maintenant, au contraire, il s'agit proprement d'un apprentissage de réception du code Morse. Après avoir pris connaissance du code, les sujets écoutent une dictée : un caractère dure $2/3$ et $4/5$ sc environ ; le temps de réponse après chaque caractère est de 3 sc, et le manipulateur du télégraphe annonce ensuite quel caractère il vient d'envoyer. Les caractères sont présentés au hasard ("carefully put into random order, dit Plotkin !) par dictées de 100 ; mais dès qu'un caractère a été appris par le sujet, (i.e. dès que le sujet donne trois fois consécutives pour ce signal x la réponse correcte ; ce qui ne signifie pas qu'il ait cessé de répondre, à tort, x pour tel autre signal y), il cesse de figurer dans les dictées.

Finalement, Plotkin range en deux tableaux les résultats cumulés de l'apprentissage des vingt sujets. Sur le premier (Plotkin p. 14), on lit chaque caractère x des données que nous noterons : $Om(x)$, nombre d'omissions (un sujet entend x , et ne met pas de réponse sur sa feuille) ; $Sub(x)$ nombre des erreurs par substitution (un sujet entend x et écrit un $y \neq x$) ; et $Cor(x)$, nombre de réponses correctes. Par exemple $Om(A) = 20$; $Sub(A) = 5$; $Cor(A) = 66$, soit au total pour les 20 sujets 91 présentations du caractère A pour que ce caractère soit appris ; tandis que pour le P , caractère beaucoup plus difficile à apprendre, il aura fallu 532 présentations, etc.. Un second tableau (Plotkin p. 16) recense les erreurs de confusion : nous noterons $k(x, y)$ le nombre de fois que le caractère x ayant été dicté, un sujet a répondu y ; Plotkin laisse vide la diagonale de ce tableau 36×36 .

On a analysé les quatre tableaux suivants :

1/ tableau k_1 , 36×36 : c'est le tableau k de Plotkin, avec $k(x, x) = 0$

2/ tableau k_2 , 36×36 : c'est le tableau k de Plotkin, où l'on a porté sur la diagonale les réponses correctes : $k_2(x, x) = Cor(x)$.

3/ tableau k_3 , 36×37 : c'est le tableau k_1 complété par une colonne supplémentaire (colonne Om) : $k_3(x, Om) = Om(x)$, correspondant à la réponse omission.

4/ tableau k_4 , 36×37 : c'est le tableau k_2 complété par la colonne Om .

Nous commenterons ici l'analyse du tableau k_2 , en y rapportant sommairement quelques particularités des autres analyses.

La suite des valeurs propres issues du tableau k_2 est : 0,73 ; 0,68 ; 0,63 ; 0,59 ; 0,55... ; et la première valeur propre ne représente que 6,83 % de la trace (inertie totale du nuage). L'analyse des tableaux k_1 et k_3 (tableaux à diagonale vide) fournit des valeurs propres plus faibles, qui épuisent cependant plus vite la trace : on a pour k_1 la suite : 0,39 ; 0,38 ; 0,24 ; 0,21 ; 0,20 ... , la première valeur propre étant 14,1 % de la trace . Ce phénomène est général (et mérite une étude algébrique détaillée, faite ailleurs, au moins pour les tableaux symétriques) : en chargeant la diagonale d'un tableau $n \times n$, on tend à obtenir un tableau diagonal, cas où l'analyse des correspondances donne $n - 1$ valeurs propres égales à 1. Mais sur le tableau à diagonale vide, la similitude entre un x et un x' ne peut se manifester que par l'existence de réponses z autres que x et x' et telles que $k(x, z)$ et $k(x', z)$ soient simultanément élevés : on perd ainsi le signe majeur de la similitude entre x et x' , signe obtenu pour $z = x$ et $z = x'$; d'où un certain désordre parfois observé dans l'analyse des tableaux à diagonale vide (bien que la théorie montre que certaines modifications de la diagonale n'influent pas sur les facteurs). Dans la présente étude, l'analyse des tableaux à diagonale vide, tout en fournissant les mêmes groupements que l'analyse des k_2 , donne des facteurs moins bien interprétables ; mais on peut voir dans la suite des valeurs propres issues de k_1 ou k_3 l'utile indice que le nombre de dimensions relevantes est peu élevé. (A la limite de l'apprentissage parfait les signaux sont parfaitement distingués : il y a, en ce sens, autant de dimensions que de caractères ! mais seule nous intéresse la phase initiale où les signaux ne sont encore pour les sujets, que les éléments indistincts d'un continuum dont quelques qualités donnent les axes.

Sur la figure 8.1 on a représenté dans le plan des deux premiers facteurs les résultats de l'analyse du tableau k_2 ; on y a joint un intéressant diagramme de proximités publié par R.N. Shepard (1964c).

Notons d'abord que bien que le tableau k_2 ne soit pas exactement symétrique, les représentations de chaque signal comme stimulus et comme réponse coïncident presque : à l'échelle de la figure 8.1, les plus grandes distances observées sont d'environ 2 cm (entre ST et RT ou S9 et R9) ; dans la zone de plus faible densité du diagramme les distances sont de l'ordre du mm avec un maximum inférieur à 0,5 cm (e.g. 3 mm entre SD et RD). Aussi pour alléger le graphique, n'a-t-on marqué que le point moyen entre stimulus et réponse.

Sur le premier axe E, T (. et -) puis I, N, A, M (signaux à deux éléments) s'opposent aux chiffres (signaux formés de cinq éléments). Sur le second axe, zéro, 5, H, S, O suites d'éléments identiques (5 traits, 5 points etc...) s'opposent à Y, C et autres signaux complexes où les traits alternent avec les points. Examinons attentivement la zone dense qui occupe le coin supérieur droit de la figure 8.1. On y voit un groupe de chiffres

{1, 2, 3, 6, 7, 8}; un groupe de signaux à quatre éléments {B, J, Z, P, Q, F, X, L, C, Y}; et les signaux à quatre éléments {U, G, D, W, K, R}. Une parallèle au 1er axe menée, e.g., par le milieu du segment DK séparerait à droite l'ensemble {K, R, P, F, Q, X, C, L, Y} des signaux les plus complexes comportant plus d'un changement du trait au point ou du point au trait : Y, par exemple, comporte deux changements -. --; C en comporte trois -. -. ; tandis que {Z, D, W, G, J, B, V, 7, 8...} ne comportent qu'un changement : par exemple Z est fait de deux traits suivis de deux points. On notera que les signaux 9, 4, V formés d'une suite d'éléments identiques puis d'un élément distinct (V...-) sont plus proches des signaux purs (H...) que ne le sont les signaux renversés (1, 6, B-...) : cette particularité nous paraît liée au fait, déjà évoqué au n° 7, que la durée du dernier élément d'une suite est perçue avec inexactitude. Cependant tous les signaux purs : zéro, 5, H, S, O, M, I, T, E se placent seuls sur une sorte de branche d'hyperbole qui est la frontière gauche du diagramme.

Quant à l'opposition point-trait, l'examen attentif du tableau d'erreurs en fait voir des vestiges (... est confondu avec ... et non avec ----), mais ni Shepard ni nous-mêmes ne l'avons vu apparaître comme un facteur ; ce qui laisse penser que l'importance en est faible. A posteriori, il est permis de proposer une explication psychologique. Dans une comparaison de paire (cf. n° 7) la durée du signal total ainsi que celle des éléments apparaissent nettement. Le signal ... entendu après le signal --- se distingue de celui-ci à la fois par la brièveté de l'ensemble et par celle des éléments ; mais après 20 sc d'intervalle le ... qu'on vient de dicter évoque le --- déjà entendu et rien ne l'en distingue d'abord. Au contraire le signal c -. -. , entendu après 0 --- ressemble à celui-ci en ce qu'il a même durée totale ; mais après un intervalle, dans une dictée, le signal C au rythme alternant n'évoque nullement la suite égale des traits du 0. On conçoit que les deux facteurs nombre d'éléments et uniformité soient prépondérants dans l'apprentissage du code.

Ces deux facteurs issus de l'analyse des correspondances ont déjà été reconnus par Shepard qui les nomme : hétérogénéité et longueur (length). Nous avons vu au n° 7 que ce dernier terme recèle une confusion : il s'agit en fait ici non de durée mais de nombre d'éléments. Quant aux groupements et à la disposition même au sein de ceux-ci, analyse des proximités (par Shepard) et analyse de correspondance s'accordent pleinement. Mais le diagramme de Shepard l'emporte sur le nôtre par une densité plus uniforme du nuage des points. A examiner le tableau d'erreurs publié par Plotkin, il apparaît, croyons-nous, que l'échelle des distances perceptives a l'irrégularité que montre le diagramme d'analyse des correspondances : un îlot tel que {1, 2, 3, 6, 7, 8} est bien plus resserré que E, T, IM). L'analyse des correspondances, partant de la matrice d'inertie du nuage rapporte tout au cadre sous-tendu par les éléments les plus éloignés. L'analyse des proximités cherche à figurer exactement le plus grand nombre possible d'inégalités entre les distances : ainsi c'est au contraire la périphérie qui est rapportée au noyau du nuage. De plus l'agorithme de Shepard qui part d'un ensemble de points quasi équidistants pour l'aplatir, nous paraît devoir tendre à effacer les pointes. Dans le cas présent c'est un avantage ; au n° 1 (cf. fig. 1.3, 1.4, 1.5), l'analyse des proximités arrondit le triangle des couleurs : c'est inopportun ; au n° 7, la figure d'ensemble issue de l'analyse des proximités est la plus harmonieuse mais un facteur très intéressant et

nettement présent - le facteur durée totale - s'estompe dans les manipulations géométriques de réduction à la dimension 2. La leçon de ces expériences pourrait se tirer ainsi : l'analyse des correspondances, soumet les données brutes à des calculs déterminés (recherches de vecteurs propres) : les résultats sont donc sûrs, sans arbitraire. L'analyse des proximités (qui demande, e.g. 20 fois plus de temps pour l'analyse de 40 points, et plus encore pour 100 points, que l'analyse des correspondances) introduit l'arbitraire du choix d'un indice de similitude et celui d'un algorithme d'optimisation : mais ces manipulations plus violentes peuvent conduire à une représentation plus satisfaisante : plus achevée sinon plus fidèle.

Signalons que l'introduction d'une colonne Omission (tableau k₄) accentue encore l'inégalité d'échelle du diagramme d'analyse des correspondances. Entre les signaux simples dont l'apprentissage est très facile et les autres signaux, s'ajoute encore cette différence que ceux-ci sont l'objet d'un très grand nombre d'omissions (e.g. pour le P, 240 omissions sur 532 présentations !) : d'où un îlot encore plus dense des signaux complexes.

La position excentrique des éléments E et T, objets de très peu de confusions, suggère de reprendre l'analyse de la matrice de confusion privée des lignes et colonnes E et T. On a ainsi analysé la matrice k₂ et cette même matrice symétrisée. Disons d'abord que (comme il fallait s'y attendre, les paires stimulus-réponse étant quasi-confondues) les résultats ne sont guères modifiés par la symétrisation. Examinons donc sur la figure 8-1' les résultats de l'analyse du tableau symétrisé. La suppression de E et T n'a elle-même que peu d'influence. La densité du nuage devient un peu plus uniforme ; les axes 1 et 2 sont échangés, mais admettent toujours la même interprétation : dans la présente analyse, le premier facteur est l'homogénéité ; le deuxième, le nombre d'éléments. Quant à la troisième valeur propre elle est un peu inférieure aux deux premières :

$$(\lambda_1 = 0,79 \text{ , } \lambda_2 = 0,77 \text{ , } \lambda_3 = 0,68) ;$$

et le 3ème facteur n'apporte pas d'information, car il semble être une fonction quadratique des deux premiers. Ajoutons au graphique les points E et T, comme des éléments supplémentaires de masse évanescence (cette adjonction n'influe pas sur la place des autres éléments ; E et T sont seulement, placés d'après leurs confusions avec les autres signaux) : E et T dans le plan rapporté aux axes 1 et 2 sont quasi-confondus ; ils n'occupent pas une position extrême, mais se placent au voisinage de M (--) et plus près de l'origine que celui-ci. (Cette position relativement centrale s'explique notamment parce que en plaçant E et T comme éléments supplémentaires, on ne tient pas compte des réponses exactes E-E, T-T, ni des confusions entre E et T)

8.2 - Considérons brièvement une expérience d'apprentissage du code Morse très semblable à celle de Plotkin : l'expérience de F.S. Keller et R.E. Taubman. Les sujets de cette expérience apprenaient le code Morse par des séances de dictées d'une heure environ où la série des 36 signaux leur était présentée 3 ou 4 fois dans son intégrité, en ordre aléatoire ; trois secondes après l'exécution d'un signal, l'opérateur en donnait la valeur exacte. L'instruction d'un sujet était jugée terminée après qu'à trois dictées successives du code entier celui-ci eût fait au total moins de 5 % d'erreurs.

Les auteurs ont tabulé pour 50 sujets l'ensemble des confusions commises dans la première série de chaque heure de dictée. Comme, selon Keller et Taubman, l'instruction d'un sujet demande, en moyenne douze dictées, on peut conclure que le tableau contient les erreurs commises sur environ $50 \times 12 = 600$ séries de 36 signaux. Ceci suggère, de remplir la diagonale du tableau d'erreurs par des nombres de réponses correctes calculés en retranchant de 600 le nombre des erreurs commises (somme des nombres de la ligne correspondant au stimulus). Cette reconstruction étant quelque peu conjecturale, on a analysé trois tableaux obtenus en supposant qu'il y avait 500, 1000 et 2000 séries : les résultats obtenus ne diffèrent pas notablement. De plus, comme au n° 8.1, on constate que stimuli et réponses tombent presque aux mêmes points ; ce que l'on a encore vérifié en analysant les tableaux symétrisés ($k_{sym}(x, y) = k(x, y) + k(y, x)$). Nous ne figurons, donc que l'analyse du tableau symétrique construit sous l'hypothèse qu'il y eut 500 séries (cf. fig. 8.2 et 8.3).

La suite des valeurs propres est : 0,97 ; 0,87 ; 0,84 ; 0,77 ; 0,75... , la première de celles-ci étant 6,87 % de la trace. Comme le premier facteur oppose seulement E et T (point . et trait -) aux 34 autres signaux nous figurons le plan des axes 2 et 3. Mis à part E et T, dont la place n'est pas significative, les signaux s'organisent suivant les mêmes facteurs qu'au n° précédent. Comme sur la figure 8.1, il y a un îlot de forte densité ; mais la répartition d'ensemble est plus nette et plus satisfaisante. Une parallèle au 3ème axe passant par le milieu du segment 6J sépare les 10 chiffres des autres signaux (E mis à part) ; le 3ème axe lui-même sépare les signaux à 5 ou 4 éléments, de ceux à 2 ou 3 éléments ; seul fait exception le H (...) qui est au dessous du 3ème axe (proche de S ...) ; comme de plus les signaux IMAN (à deux éléments) sont quasi-alignés sur une parallèle au troisième axe, on voit que l'interprétation du 3ème facteur comme nombre d'éléments est quasi-parfaite. Quant à l'hétérogénéité, déjà vue sur la figure 8.1, elle apparaît en deuxième facteur, les signaux comportant plus d'une alternance (de point à trait ou de trait à point) sont nettement groupés ; suivent les signaux à une alternance ; puis les signaux purs auxquels se mêlent 4 et 9.

Ici encore, une comparaison avec l'analyse des proximités s'impose. Pour les raisons déjà données, l'analyse des correspondances isole E et T ; et le diagramme a des plages désertiques ... L'analyse des proximités évite cela mais elle n'inscrit pas à son palmarès un 2ème et un 3ème facteur nettement interprétables ...

Notons les résultats de l'analyse du tableau symétrisé à diagonale vide. On retrouve les remarques faites au n° 8.1. La trace est plus faible qu'avec diagonale pleine ; les valeurs sont propres 0,69 ; 0,65 ; 0,56 ; 0,51 ; 0,27... ; la première étant 13 % de la trace. Les deux premiers facteurs servent à isoler E (qui ne se confond notablement qu'avec T) et T (qui ne se confond notablement qu'avec E) : ici il n'y a pas une classe {E, T} mais deux classes {E}, {T}, puis les 34 symboles restants. Dans le plan des facteurs 3, 4 on retrouve les groupements de la figure 8.2 : La disposition d'ensemble est moins satisfaisante ; mais les classes sont nettes et la brusque chute entre λ_4 et λ_5 - de 0,51 à 0,27 - accroît la confiance en une représentation bidimensionnelle.

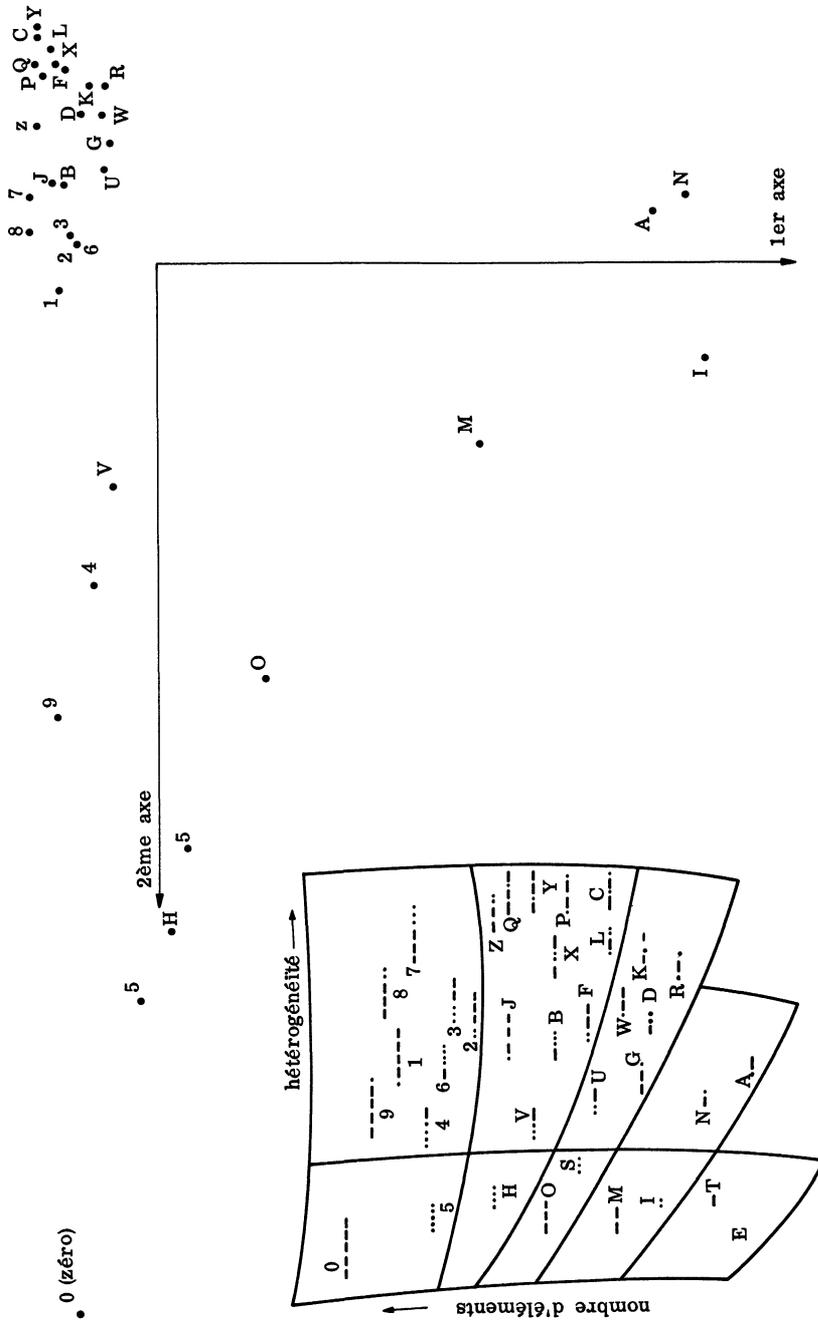


Figure 8.1 - Apprentissage de Code Morse d'après L. Plotkin. Sur le diagramme d'analyse de correspondance, on a donné aux axes la longueur 2. En cartouche, on a figuré un diagramme d'analyse de proximité, d'après R.N. Shepard (1964 c).

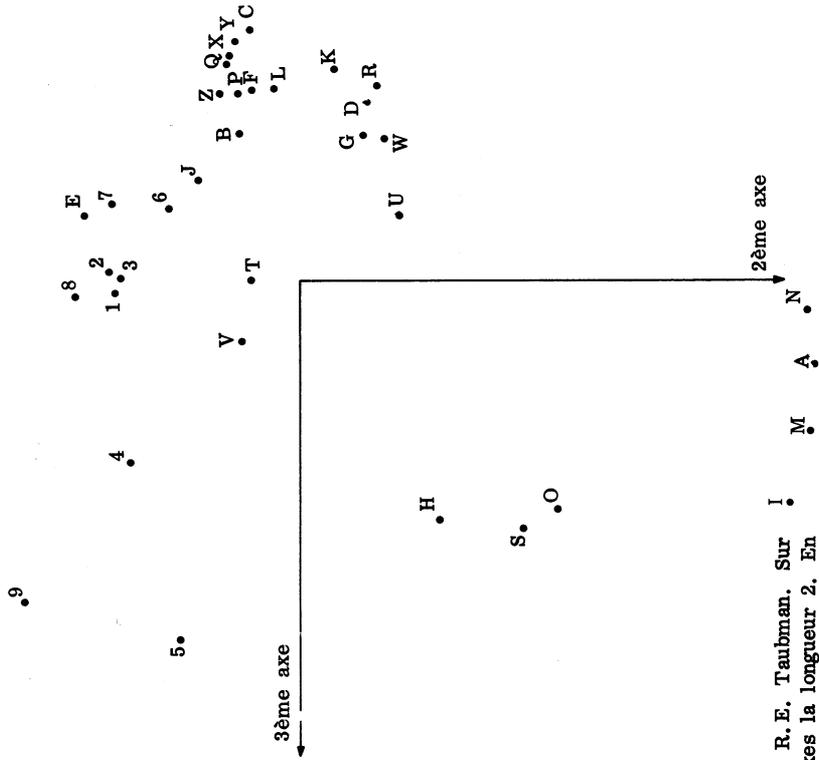
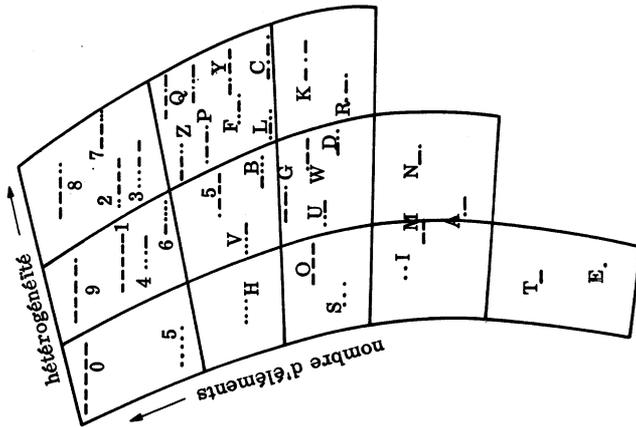


Figure 8.2 - Apprentissage du Code Morse selon F.S. Keller et R.E. Taubman. Sur le diagramme d'analyse des correspondances, on a donné aux axes la longueur 2. En cartouche, on a figuré le diagramme d'analyse des proximités, d'après R.N. Shepard (1964, c).

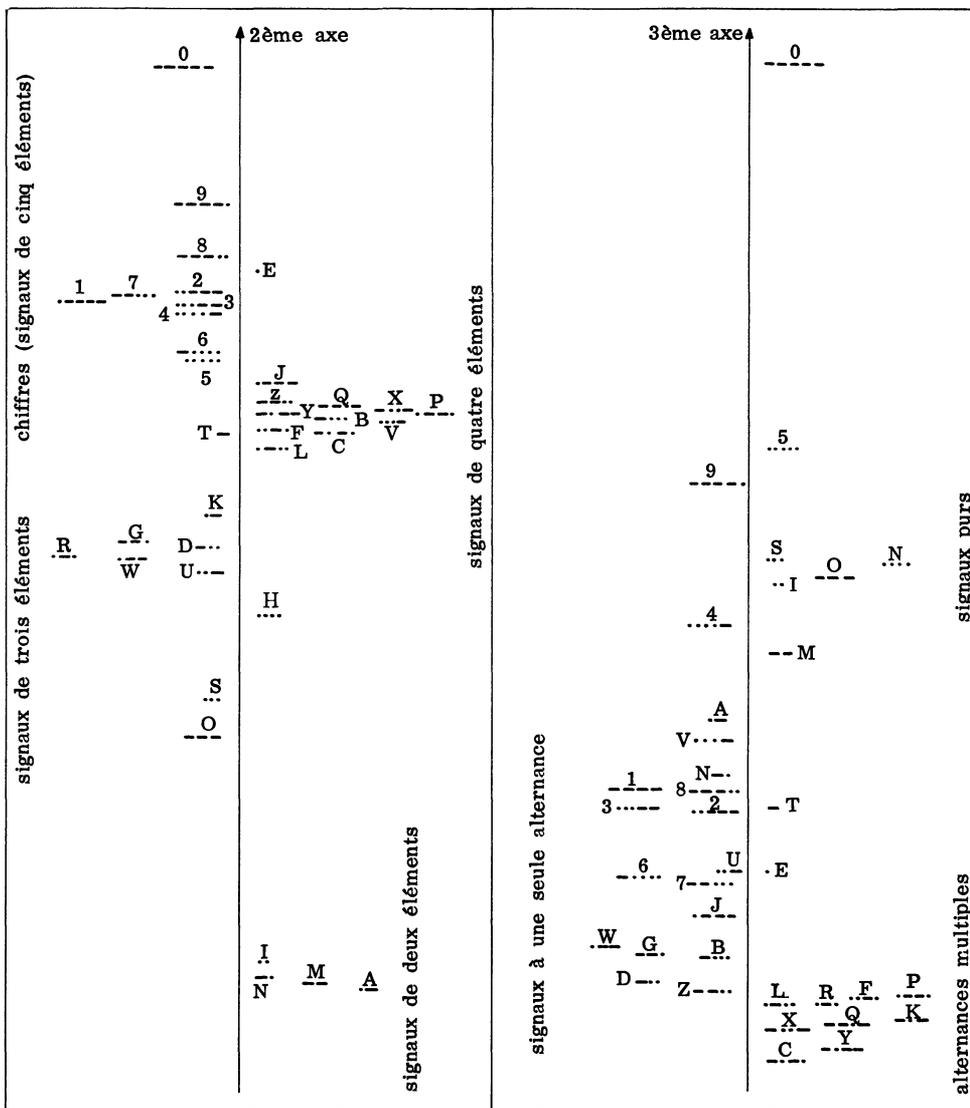


Figure 8.3 - Apprentissage du code Morse d'après Keller et Taubman. A droite, diagramme du 3ème axe ; à gauche, diagramme du 2ème axe. Sur chaque axe, on a distingué les groupes de signaux placés algèrnativement à droite et à gauche et correspondant à l'interprétation du facteur : nombre d'éléments pour le 2ème et complexité (nombre d'alternances de point à trait ou de trait à point) pour le 3ème. Les signaux E et T, séparés par le 1er facteur, n'ont pas de position significative.

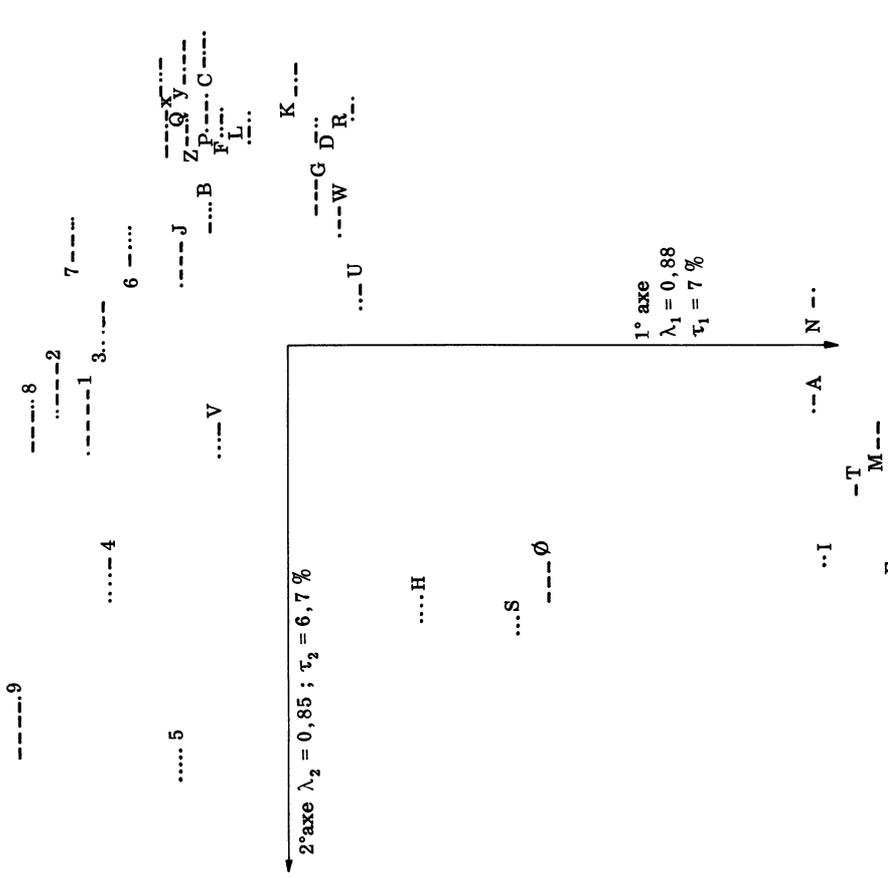


Figure 8.2' - Apprentissage de Code Morse d'après F.S. Keller et R.E. Taubman. On a donné aux axes longueur 2. Les signaux E et T sont traités comme éléments de masse nulle.

Plus encore qu'au n° 8.1, la place singulière des signaux E et T, invite à reprendre l'analyse de la matrice de confusion privée des lignes et colonnes E et T. On a analysé le tableau carré complété sous l'hypothèse qu'il y eut 500 séries. Les paires stimulus-réponse étant quasi-confondues on s'est borné, sur la figure 8.2', à représenter chaque symbole une fois. On voit que les facteurs 1 et 2 issus de la présente analyse ne sont autres que les facteurs 2 et 3 issus de l'analyse du tableau complet (comportait lignes et colonnes E et T). Comme éléments supplémentaires, E et T (signaux à un élément) se placent avec I, M, A, N (signaux à deux éléments) et du côté de I, M (signaux purs) : ce qui est satisfaisant. Quant à la 3ème valeur propre elle est un peu inférieure aux deux premières

$$(\lambda_1 = 0,88, \lambda_2 = 0,85, \lambda_3 = 0,78) ;$$

et le 3ème n'apporte pas d'information car il semble être fonction des deux premiers.

9 - RECONNAISSANCE DES FORMES

Dans le domaine de la reconnaissance des formes, n'ayant pas encore analysé de données importantes nous nous bornerons à suggérer des expériences.

9.1 - Confusions entre visages : on choisit un certain nombre de photographies de visages, par exemple 50. L'une de ces photographies, i , est montrée, à un assez grand nombre de sujets en vision brève ; puis on fait passer la série des 50 photographies en avertissant les sujets qu'ils doivent à toute force reconnaître la personne qui a été vue d'abord, en désignant au besoin plusieurs des 50 personnes qui leur sont présentées à la suite (J.F. Richard nous a suggéré d'annoncer l'expérience comme une recherche de coupable) : on note $k(i, j)$ le nombre de sujets qui lors de la présentation de l'image j , ont dit y reconnaître l'image i qui leur avait été montrée d'abord. Ainsi quand toutes les images ont été successivement présentées en tête de série, on a rempli, ligne après ligne, une matrice de confusions. Grâce à la consigne donnée aux sujets de reconnaître à toute force, au risque de faire des erreurs (de prendre des innocents plutôt que de laisser fuir le coupable), la matrice doit être remplie avec une densité suffisante. Toutefois certaines difficultés se présentent. Il ne faut pas que les sujets soient familiarisés avec l'ensemble des visages, car alors ils ne feront guère d'erreurs et attendent le passage de l'image entrevue d'abord ; aussi semble-t-il nécessaire de ne pas remplir toutes les lignes de la matrice en interrogeant le même groupe d'observateurs ; ce qui complique l'expérimentation. D'autre part, fût-ce au sein d'une série unique, les sujets, une fois reconnu le visage même qu'ils ont entrevu d'abord, ne seront plus enclins à croire le reconnaître dans les images qui sont présentées après lui. On peut obvier à cela en ne présentant pas avec la série, ou en présentant le dernier, le visage qu'il s'agit de reconnaître : cet artifice n'échapperait peut-être pas aux sujets, s'ils voyaient 50 séries, mais nous avons dit qu'il ne semble pas possible de recourir au même public pour remplir toutes les lignes de la matrice de confusion.

Pour rendre la tâche de reconnaissance assez difficile et amener ainsi le taux de confusion au niveau le plus favorable à l'analyse de la matrice,

on a proposé ci-dessus de ne présenter que brièvement le visage i à reconnaître. On pourrait procéder autrement. Suggérons quelques variantes. Le visage i est présenté de face, tandis que la série des images ne comporte que des photographies prises de profil. Ou encore le visage i est en partie caché : on ne voit que la partie inférieure du visage etc... Ou encore, le visage à reconnaître est seulement décrit par un signallement qui est lu aux sujets.

L'expérience pourrait être grandement simplifiée par un système d'enregistrement automatique des données : les sujets n'auraient qu'à presser un bouton pour annoncer qu'ils ont cru reconnaître l'image première et le calcul (par comptage des boutons pressés) de $k(i, j)$ serait fait automatiquement.

9.2 - Peintres et tableaux : Nous avons d'abord considéré comme ensemble de formes un ensemble de visages : l'analyse des matrices de confusion nous paraît mériter d'être tentée dans bien d'autres domaines. Prenons pour exemple la peinture. Il ne semble pas très intéressant de présenter un tableau - disons la Vierge aux rochers de Vinci - puis 50 tableaux parmi lesquels il faudrait reconnaître le premier présenté : on n'aurait guère que des confusions portant sur le sujet (avec une autre image de la Vierge...), à moins de n'avoir choisi des tableaux tous très voisins de thèmes... Mais on peut présenter d'abord un portrait, puis 50 paysages, parmi lesquels il faut reconnaître celui qui a été peint par le même auteur que le portrait. Ou encore présenter un détail très agrandi faisant voir la touche d'un artiste, puis 50 autres détails dus à divers peintres. Ou encore présenter d'abord un tableau, puis 50 détails etc... Comme il s'agit ici d'une tâche assez difficile, il n'y aurait pas d'inconvénient à laisser les documents à la disposition des sujets qui rempliraient à loisir un questionnaire.

9.3 - Associations complémentaires : Les expériences proposées aux n° 9.1, 9.2, comme toutes celles analysées plus haut, concernent des confusions ou, tout au moins des associations par affinités par similitudes. On aboutit à un tableau carré $k(i, j)$, à peu près symétrique ; et $k(i, j)$ est d'autant plus grand que i est plus voisin de j . Par exemple si i est un tableau et j un détail agrandi, $k(i, j)$ sera d'autant plus grand que le peintre, dont le détail j nous révèle la touche, a une manière plus semblable à celle de l'auteur du tableau i ; et finalement l'on aura rempli un tableau de confusion entre peintres.

Cependant l'on sait que l'analyse d'un tableau carré $I \times J$ symétrique ($k(i, i') = k(i', i)$) peut conduire à des facteurs soit directs soit inverses : on parle de facteur direct si $F(i) = G(i)$ et de facteur inverse si $F(i) = -G(i)$. Ce dernier cas correspond à une dimension d' , associations par anti-thèse, ou tout au moins par complémentarité. Eventuellement, une même analyse révélera à la fois des facteurs directs et des facteurs inverses.

Soit une série d'objets différents quant au style et à la couleur : par exemple des carrés de tissu, ou des vêtements. On demande aux sujets, étant donné un objet i , d'en choisir un autre i' qui s'assortisse avec i : et l'on compte $k(i, i')$, nombre de sujets qui ont choisi i' pour aller avec i . Certains facteurs de style, de matière joueront dans le sens direct : on mettra ensemble des objets de styles voisins etc... Mais la couleur produira des associations par antithèse, par complémentarité. Afin d'avoir les

données simples, où ne joue qu'un seul facteur, on pourrait prendre pour ensemble d'objets le catalogue des échantillons de couleurs d'une marque de laine à tricoter ou d'étoffe...

BIBLIOGRAPHIE

- N.B. - Les publications multigraphiées marquées I.S.U.P., se trouvent au laboratoire de Statistique de la Faculté des Sciences de Paris.
- J.P. BENZECRI - Sur l'instauration d'un code ; Rennes (1965) et I.S.U.P.
- J.P. BENZECRI - Analyse statistique et modèles probabilistes en psychologie ; in Revue de l'Institut international de statistiques ; vol 34 n° 2 pp. 139-155 (1966) (a).
- J.P. BENZECRI - Arbitrarité et analogie dans l'instauration d'un code ; in Colloque d'informatique, Toulouse (1966) (b) et ISUP.
- J.P. BENZECRI - Statistical analysis as a tool to make patterns emerge from data ; in Proceedings of the 1968 Honolulu conference on pattern recognition, ed. by S. Watanabe ; Acad. Press (1969).
- G. EKMAN - Dimensions of color vision ; in The Journal of Psychology, Vol 38, pp. 467-474 ; (1954).
- S.C. JOHNSON - Hierarchical clustering schemes ; in Psychometrika Vol 12, n° 3 ; pp. 241-254 ; (1967).
- F.S. KELLER & R.J. TAUBMAN - Studies in international Morse code 2. Errors made in code reception ; in Journal of experimental psychology ; Vol 40 ; pp. 504-509 ; 1943.
- G.A. MILLER & P.E. NICELY - An analysis of perceptual confusions among some english consonants ; in Journal of the Acoustical Society of America ; Vol 27 n° 2 pp. 338-352 ; (1955).
- F. PARRA - Recherches sur le seuil différentiel de couleur ; Thèse, Paris ; (1966).
- L. PLOTKIN - Stimulus generalization in Morse code learning ; in Archives of Psychology N.Y. Vol 40 n° 287, (1943).
- J.F. RICHARD & ROUANET - Analyse d'une situation de détection et de localisation ; in l'Année psychologique pp. 37-68 ; (1968).
- E.Z. ROTHKOPF - A measure of stimulus similarity and errors in some paired-associate learning-tasks ; in Journal of Experimental Psychology, Vol 53, n° 2 pp. 94-101 ; (1957).
- R.N. SHEPARD - The analysis of proximities : Multidimensional scaling with an unknown distance function ; Psychometrika, Vol 27 pp. 125-140 & 219-246 (1962).
- R.N. SHEPARD - Analysis of Proximities as a Technik for the Study of Information Processing in Man ; in Human Factors pp. 33-48 ; (1963).

- R. N. SHEPARD - Attention and the Metric Structure of the Stimulus Space; in Journal of Mathematical Psychology ; Vol 1 n° 1, pp. 54-87 (1964); (a)
- R. N. SHEPARD - Circularity in Judgments of relative pitch ; in The Journal of the Acoustical Society of America ; Vol 36 n° 12, pp. 2346-2353 (1964) ; (b).
- R. N. SHEPARD - Extracting latent structure from behavioral data ; in Proceeding of the 1964 symposium on digital computing January 30-31, 1964, at Bell Telephone Laboratories, Holmdel (1964) ; (c).
- P. STREVEENS - Spectra of fricative noise in human speech ; in Language and Speech, Vol 3, pp. 32-49, (1960).
- K. V. Wilson - Multidimensional Analysis of Confusions of English consonants ; in ... ? pp. 89-95 (1960) ou 1961).