

## BILINGUISME ET TIMBRE VOCAL\*

*par Alan Cranshaw*

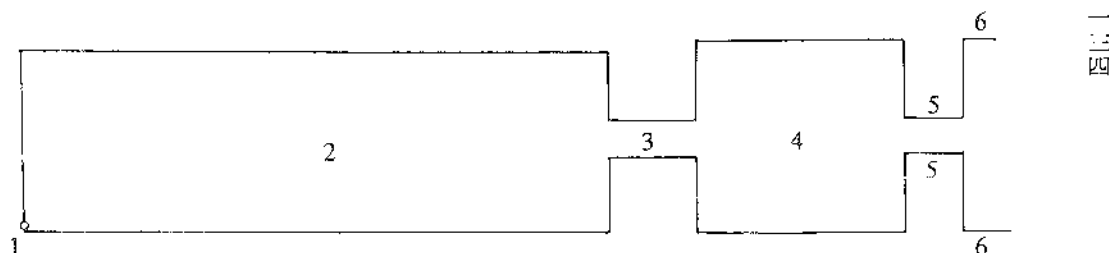
### Introduction et Généralités

L'identification vocale fait partie de notre vie quotidienne. Tout le monde est capable de reconnaître la voix de ses amis, et l'on ne se trompe que rarement. Le timbre de la voix d'un individu est quelque chose d'indiscutablement personnel. Des expériences suggèrent, en outre, que le timbre de la voix est dépendant de la langue qu'on parle.<sup>1</sup> Si cela est vrai, les bilingues ont-ils la capacité de moduler le timbre de leur voix selon la langue qu'ils parlent. Autrement dit, est-ce qu'un bilingue change la qualité de sa voix quand il passe d'une de ses langues à l'autre.

Le son qu'on appelle la voix est le résultat d'une vibration des cordes vocales. Celles-ci se trouvent à l'intérieur du larynx, qui se situe dans la trachée. La capacité de mouvement des cordes vocales, et la maîtrise de notre appareil phonatoire nous permettent de produire des sons vocaux.

Bien que la vibration des cordes vocales ne crée que la fréquence fondamentale de la voix, la voix est plus qu'une fréquence fondamentale. En effet, il y a certaines résonances bien déterminées qui s'ajoutent à la fréquence fondamentale pour former notre voix. Ces résonances sont engendrées dans nos cavités buccale, nasale, et pharyngale. Si, par exemple, on place une source sonore au milieu d'une grande pièce, certains harmoniques sonneront plus que d'autres à cause des caractéristiques de résonateur de la pièce en question. Nos cavités buccale, nasale, et pharyngale servent ainsi de résonateurs avec une différence très importante, à savoir, la variabilité de la taille des cavités. Ce fait nous permet de modifier à notre gré les harmoniques qui résonneront dans notre voix. (La cavité nasale n'étant pas de taille variable est donc un cas à part — elle ne peut être qu'ouverte ou fermée).

Le schéma suivant est une représentation des cavités buccale, nasale, et pharyngale ou se trouvent: 1. l'emplacement des cordes vocales, 2. la cavité pharyngale, 3. le point de rapprochement le plus important entre la langue et le voile du palais ou paroi pharyngale, 4. la cavité buccale, 5. les dents et 6. les lèvres.



Quand nous déplaçons notre langue et ainsi changeons le point 3., la taille des deux cavités 2. et 4. change, et avec elle, les fréquences renforcées dans chacun de ces résona-

teurs. Ce jeu de résonateurs couplés est responsable en grande partie du timbre des voyelles que nous produisons, mais il a aussi un rôle relativement important dans la qualité de notre voix en général. Une voix est, en fait, la somme d'un grand nombre de traits acoustiques qui se combinent pour produire un effet sonore. Cet effet est probablement unique, et plusieurs facteurs ont été postulés comme responsables de sa qualité.

La structure du corps, selon Schilling,<sup>2</sup> influe sur la fréquence fondamentale, ainsi que la taille des cordes vocales. La forme et la taille des cavités résonatrices jouent un rôle important dans la qualité de la voix aussi, surtout celles de la cavité pharyngale.<sup>3</sup>

D'autres facteurs qui peuvent contribuer à la qualité de la voix sont la façon de respirer et la vigueur de respiration, l'état thymique, et le contenu du message que l'on veut communiquer. Bien qu'importants dans l'effet sonore produit, ces facteurs sont d'un intérêt secondaire car tous varient beaucoup.

La qualité de la voix est donc quelque chose de plus ou moins déterminé par la physiologie du corps. Pour utiliser une métaphore, la qualité de la voix est comparable au timbre d'un instrument musical (on parle aussi du "timbre" de la voix). En quelque sorte, on "joue" de la voix pour fabriquer le langage tout comme on joue d'un instrument pour faire de la musique. La voix n'est pas le but d'un énoncé, mais elle est le médium dont on se sert pour rendre audible le langage. Comment alors peut-on décrire le timbre de la voix de façon précise? En fait, le timbre de la voix et le langage sont deux choses inextricablement mêlées car c'est à travers le langage que notre voix possède un timbre. Le timbre d'un son peut être défini par "le nombre et les intensités relatives des harmoniques qui le composent".<sup>4</sup> Toutefois, cette définition doit être modifiée pour pouvoir s'appliquer de façon adéquate à la voix, car à un moment donnée, toutes les fréquences qui composent notre voix ne sont pas représentées. Les fréquences produites à n'importe quel instant par notre voix sont une fonction de ce que nous sommes en train de prononcer. Le timbre de la voix doit être défini comme "l'ensemble des fréquences les plus usuelles, et leurs intensités relatives qui se combinent pour caractériser la voix".

La distinction entre timbre vocal et accent est très importante — l'accent est la façon dont on se sert de la voix pour émettre des messages linguistiques tandis que le timbre de la voix est le son vocal qu'un locuteur produit afin de le façonner ensuite pour former l'accent. Alors que l'accent est généralement considéré comme un phénomène national ou régional, le timbre de la voix est strictement personnel. Autrement dit, l'accent nous divulgue l'origine géographique ou sociale d'un locuteur en question tandis que le timbre de sa voix nous permet de l'identifier individuellement.

Le but de cette recherche est de savoir si un bilingue change la qualité de sa voix quand il change de langue. Pour ce faire, j'ai été amené à étudier les techniques de l'identification vocale. C'est à dire, comment est-ce que l'on peut identifier avec plus ou moins de certitude, un échantillon de parole enregistrée comme étant la voix d'une personne déterminée.

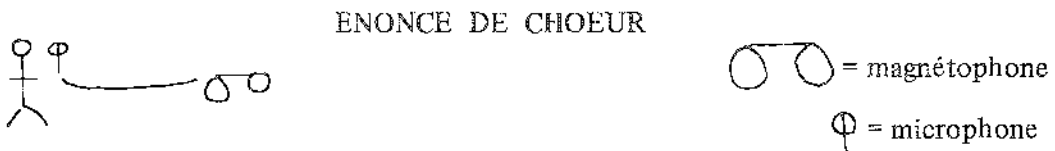
Car, s'il est vrai qu'un bilingue change la qualité de sa voix quand il change de langue, alors les tests d'identification vocale doivent systématiquement échouer quand l'on essaie d'identifier les deux langues d'un bilingue comme étant d'une seule et même voix.

## L'Identification Vocale

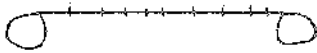
Regardons brièvement les sortes de tests d'identification vocale utilisées dans cette expérience.

1. L'identification par test auditif. Un groupe de personnes (auditeurs) entend deux voix enregistrées et doit décider si les deux voix émanent d'une seule personne ou non.

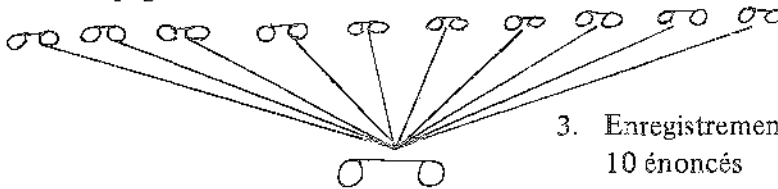
2. L'identification par spectres de chœur. Un spectre de chœur est un sonagramme ('sound spectrogram') fait à partir d'un énoncé de chœur. L'énoncé de chœur est une technique développée par Tarnoczy<sup>6</sup> qui facilite l'identification d'une voix. Elle consiste à enregistrer un énoncé assez long, soit de cinq minutes. Ensuite on découpe cet énoncé linéairement en au moins dix morceaux — ce qui donnerait dix énoncés de trente secondes, tous prononcés par la même personne, mais de contenu différent. Après on passe ces dix énoncés simultanément pour enregistrer 'l'énoncé de chœur'. Le résultat est un bruit qui ressemble à une foule de personnes qui parlent en même temps, à cette exception près — toutes 'les voix' que l'on entend émanent d'une seule et même personne. (voir schéma ci-dessous)



1. Enregistrement d'un énoncé de 5 minutes



2. Découpage de cet énoncé en 10 morceaux



3. Enregistrement simultané des  
10 énoncés

Cette technique permet de faire une évaluation plus précise du timbre de la voix de quelqu'un, car l'influence de la parole sur la qualité de la voix est réduite à un minimum. Avec ces énoncés de chœur on fait des spectres de chœur, c'est à dire, un graphique représentant les fréquences de la voix en question en Hz (en abscisse) et l'intensité relative de ces fréquences (en ordonnée) à un moment donné. En comparant deux spectres de chœur, on peut prendre une décision assez fiable sur l'origine des deux voix.

Les tests d'identification vocale se laissent classer, de par leur nature, sous trois rubriques: les tests discriminatoires ('discrimination tests'), les tests ouverts ('open tests') et les tests clos ('closed tests'). Dans les tests discriminatoires, on donne à l'auditeur (celui qui essaye d'identifier la voix en question) des échantillons d'une voix identifiée et d'une voix non identifiée. Il décide alors si les deux voix sont d'une seule et même personne ou non. Avec ce genre de test, deux types d'erreur peuvent se

produire — a) la fausse élimination (les deux voix sont de la même personne mais l'auditeur croit entendre deux personnes différentes) et b) la fausse identification (l'auditeur croit entendre une même voix deux fois, mais en réalité il y a deux voix différentes).

Les tests ouverts sont composés d'une voix identifiée et de plusieurs voix non identifiées. On informe l'auditeur que la voix identifiée ne se trouve pas forcément parmi les voix non identifiées. Il y a donc trois types d'erreurs possibles: a) la fausse élimination et deux types de fausse identification (suivant que la voix non identifiée se trouve parmi les voix identifiées ou non).

Dans les tests clos, l'auditeur entend une voix identifiée et plusieurs voix non identifiées, et on lui précise que la voix identifiée se trouve parmi les non identifiées. Il n'y a donc plus qu'un type d'erreur possible — la fausse élimination.

Quand ces tests d'identification vocale sont utilisés pour des identifications intralinguistiques, le taux d'identifications correctes pour les tests auditifs tourne autour de 90%.<sup>7</sup> C'est à dire que vous enregistrez une personne qui raconte une histoire dans une langue A. Vous refaites un enregistrement de la même personne qui raconte une autre histoire dans la même langue A (donc intralinguistique). En écoutant les deux enregistrements, 90% des gens sauraient vous dire s'il s'agit d'un même locuteur.

S'il est vrai que les bilingues changent la qualité de leur voix quand ils changent de langue, le taux d'identifications correctes doit être inférieur à 90%. Autrement dit, vous enregistrez un bilingue qui raconte une histoire dans une langue A. Vous refaites un enregistrement de la même personne qui raconte une autre histoire dans sa *deuxième langue B* (donc interlinguistique). Le taux d'identifications correctes doit être inférieur à 90% si cette hypothèse est correcte.

### Tests Auditifs

Cette expérience a consisté à enregistrer sept bilingues dans leurs deux langues. (Ici, le mot 'bilingue' signifie une personne qui passe pour un natif dans deux environnements linguistiques différents.) en discutant avec ces bilingues, tous m'ont affirmé que leur voix était différente dans chacune de leurs langues. De ces impressions est née cette étude, qui essaie d'établir de façon plus scientifique si un bilingue possède en fait deux timbres vocaux associés respectivement à ses deux langues.

Pour cette expérience, j'ai procédé à l'enregistrement de sept bilingues. Ils ont tous lu un texte (en général tiré d'un journal ou d'un livre au hasard) dans une de leurs langues, et raconté une histoire dans l'autre.

- |                |               |                         |
|----------------|---------------|-------------------------|
| I<br>II<br>III | bilingue n° 1 | arabe standard/français |
|                | 2             | arabe égyptien/français |
|                | 3             | alsacien/français       |
|                | 4             | breton/français         |
|                | 5             | américain/norvégien     |
|                | 6             | wolof/français          |
|                | 7             | italien/français        |

Ces 14 enregistrements de deux minutes ont servi pour les tests auditifs d'identification vocale ainsi que pour faire les énoncés de choeur.

La technique des énoncés de choeur a déjà été décrite, mais il me semble nécessaire de faire une description détaillée de la méthode qui m'a été propre pour obtenir des

énoncés de chœur pour mes sept sujets. Les 14 enregistrements effectués sur sept sujets (deux minutes en deux langues chacun) sont devenus 14 bandes "mères". Suivons de près sur un exemple de la genèse d'un énoncé de chœur: celui du locuteur n°1 en français.

Quatre copies sont faites de la bande mère; ces quatre copies sont ensuite placées sur quatre magnétophones différents reliés à une table de mixage qui, elle, est branchée sur un cinquième magnétophone. Celui-ci se met à enregistrer alors que les quatre copies sont passées en même temps avec deux secondes de décalage au départ de chaque copie. Autrement dit, la première copie est lancée sur "écoute"; deux secondes plus tard la deuxième copie commence, suivie de la troisième à deux secondes d'intervalle, puis de la quatrième, toujours avec ces deux secondes de décalage. Tout ceci est enregistré sur le cinquième magnétophone et fournit donc unregistrement du locuteur n° 1 répétant quatre fois la même chose toutes les deux secondes. On procède à cette opération trois fois, pour obtenir trois de ces bandes du locuteur n° 1 ou se chevauche un même discours à quatre moments différents. Ces trois bandes sont ensuite mises sur trois des quatre magnétophones reliés à la table de mixage. Cette fois-ci quand on lance la première bande, on attend huit secondes avant de lancer la deuxième, car sur cette première bande, le locuteur numéro 1 commence son discours à la seconde 0, et puis à la deuxième, à la quatrième, et à la sixième seconde. Grâce à la deuxième bande, le locuteur numéro 1 recommencera à la 8<sup>e</sup>, 10<sup>e</sup>, 12<sup>e</sup>, et 14<sup>e</sup> seconde, et grâce à la troisième bande à la 16<sup>e</sup>, 18<sup>e</sup>, 20<sup>e</sup>, et 22<sup>e</sup> seconde. Donc à partir de la 22<sup>e</sup> seconde on peut enregistrer le locuteur numéro 1 prononçant simultanément 12 énoncés différents. L'enregistrement ainsi obtenu, ou "énoncé de chœur", ressemble au bruit produit par une foule de gens qui parlent tous en même temps, mais tous les sons enregistrés émanent de la même personne.

Quant aux spectres de chœur, je les ai obtenus de la façon suivante: Pour chaque énoncé de chœur j'ai fait tirer 6 sonagrammes de section. Ces sections montrent l'intensité par rapport à la fréquence à un instant précis dans le temps pour les fréquences de 80 à 8000 Hz. L'utilité des énoncés de chœur dans ce procédé est que chaque énoncé donne un bon échantillonnage des fréquences les plus usuelles d'une voix à tous les moments de l'enregistrement, car celui-ci est en fait 12 enregistrements de la même personne à des moments différents.

Chaque sonagramme de section représente alors les intensités des fréquences émises par le locuteur à 12 moments différents. Ayant produit 6 sections par locuteur et par langue (6 \* 7 bilingues \* 2 langues chacun = 84 sections), j'ai obtenu une représentation des intensités des fréquences émises par chaque locuteur à 72 moments différents. Ces sections ont été transcrites sur papier millimétré transparent, et l'intensité notée à chaque millimètre (les millimètres correspondant à un écart de 80 Hz). Ensuite j'ai calculé pour chaque millimètre de graphique la moyenne des intensités des six courbes pour chaque locuteur dans ses deux langues. Le graphique ainsi obtenu est donc la section moyenne et représente l'intensité moyenne des fréquences entre 80 et 8000 Hz (par degré de 80 Hz) pour chaque locuteur à 72 moments différents. Cette méthode de spectres de chœur a été créée et utilisée avec succès par Tarnoczy, et reprise par Bordone et Tosi.<sup>8</sup>

Avant d'exposer les modalités et résultats de cette expérience, regardons quelques arguments purement théoriques pour venir à l'appui de la supposition qu'un bilingue

possède deux voix.

La perception d'une voix, comme de tout autre son, s'effectue grâce à l'effet produit par une onde sonore sur le tympan. Or, une onde sonore est une combinaison de fréquences qui possèdent, chacune, une durée et une intensité relatives. Pour que nous puissions reconnaître les phonèmes d'une langue, ceux-ci sont obligatoirement caractérisés par des fréquences spécifiques, ainsi que par des durées et intensités relatives. Puisque le stock de phonèmes, de même que la façon de prononcer les "mêmes" phonèmes varient d'une langue à l'autre, les fréquences caractéristiques de la parole changent légèrement d'une langue à l'autre. Si l'on admet que le timbre vocal est l'ensemble de fréquences les plus usuelles et leurs intensités relatives qui caractérise la voix individuelle, celle-ci change obligatoirement avec tout changement dans le stock de phonèmes, donc avec la langue. Bien que la démarcation entre fréquences caractéristiques de la langue, et fréquences caractéristiques d'une voix soit très difficile à établir, celle-ci existe néanmoins et explique la variété des timbres vocaux dans une même langue.

Le timbre vocal d'une personne intervient, entre autres, dans le choix et la bande de fréquences utilisées pour produire les sons d'une langue. Ceux-ci ne sont pas composés par un ensemble de fréquences exactes, mais varient avec chaque locuteur, et donc avec sa voix. Il est très possible alors que ces fréquences ou bandes de fréquences varient selon la langue que le locuteur bilingue se met à employer.

On peut également postuler l'existence de fréquences ou de bandes de fréquences dans la voix qui soient complètement superflues pour la parole, et qui ne font donc que contribuer au timbre de la voix. Ce serait le cas des formants au-dessus du  $f_3$ . En tout cas, les formants au-dessus du  $f_3$  ont été très peu étudiés et leur rôle dans la parole et le timbre vocal reste assez flou.

### Tests Auditifs

Pour faire cette expérience j'ai procédé de la façon suivante. Sur une seule cassette j'ai mis dix huit paires de voix; il s'agissait alors de tests discriminatoires. Chaque paire de locuteurs était précédée par un numéro d'identité. Après ce numéro, lu à haute voix, un premier locuteur parle pendant 30 secondes. Tout de suite après (sans qu'il y ait répétition du numéro d'identité) un deuxième locuteur parle pendant 30 secondes dans une langue différente de la première, après quoi on passe à la paire de locuteurs suivante (précédée d'un numéro d'identité). Les auditeurs avaient donc simplement à décider si, dans chaque paire de locuteurs, les énoncés émanaient d'une seule personne, ou de deux personnes différentes.

Les résultats de cette expérience ont été interprétés suivant 2 calculs différents: le pourcentage d'identifications correctes (le nombre d'identifications correctes divisé par le nombre total d'identifications), et le pourcentage de fausses éliminations (le nombre de fausses éliminations divisé par le nombre total d'éliminations).

Le taux d'identifications correctes pour un premier groupe de 15 auditeurs a été de 60%, pour un deuxième groupe de 39 auditeurs de 62%, et pour un troisième groupe de 64 auditeurs de 58%. Ces chiffres sont nettement plus bas que le taux moyen d'identifications correctes intralinguistiques (90%)

Pour les mêmes trois groupes d'auditeurs, les taux de fausses éliminations ont été

de 27%, 22%, et 22% respectivement. Le taux de fausses éliminations dans les tests intralinguistiques varie entre 8% et 12%.

J'estime que ces résultats sont assez clairs pour conclure que le timbre vocal d'un bilingue change avec sa deuxième langue.

J'ajoute en passant qu'un troisième calcul a été fait -- celui du pourcentage d'identifications correctes par locuteur (pour chacun des 7 locuteurs, le pourcentage d'auditeurs ayant correctement associé ses deux 'voix'). Ces calculs ne font que confirmer que certaines voix sont plus faciles à identifier que d'autres.

#### Identifications correctes par locuteur

bilingue numéro	1	83%
	2	86%
	3	25%
	4	36%
	5	97%
	6	85%
	7	81%

En outre, 39 auditeurs ont participé à un test d'identification par énoncés de chœur (taux d'identification correctes 72%, fausses éliminations 59%), mais comme ce test était seulement un test d'identification interlinguistique et qu'un test intralinguistique homologue n'a pas été fait, il est impossible d'en tirer des conclusions valables.

### Tests par Spectre de Chœur

Regardons maintenant les spectres de chœur. Pour les deux langues des sept bilingues, six spectres ont été faits, et c'est avec les six spectres de chœur que j'ai calculé le "spectre moyen" pour chaque locuteur et langue. En ce qui concerne les spectres moyens (ainsi que tous les autres utilisés dans cette expérience), il ne faut pas oublier que ces graphiques nous montrent des *intensités relatives*. L'important alors n'est pas d'avoir des intensités des fréquences exactement identiques pour tous les spectres d'une personne, mais plutôt que les pics et les creux d'intensité de deux spectres représentant un même personne soient aux mêmes fréquences. En effet, un pic désigne une fréquence très usuelle, et un creux signale une fréquence rare de la voix en question.

Une fois les spectres moyens tracés, je me suis heurté au problème de l'interprétation de ces courbes. Car les spectres de chœur ont deux grands avantages sur les autres méthodes employées jusqu'ici, à savoir, l'élimination presque totale des effets de la parole (grâce aux 72 moments ramenés à un seul instant dans le temps), et l'objectivité. Si je me servais uniquement d'une simple impression visuelle pour interpréter ces courbes, l'objectivité recherchée dans cette méthode en souffrirait. C'est pourquoi j'ai établi une échelle de "points de similitude". Cette échelle est, certes, arbitraire, et son application à ces graphiques ne peut échapper à la subjectivité, toutefois elle m'a fourni une base concrète à laquelle je pouvais me tenir pour l'interprétation des spectres. La voici:

#### *Echelle des points de similitude pour l'interprétation des spectres de chœur.*

Pics ou creux qui coïncident exactement = 10 points

Pics ou creux qui coïncident de très près = 0-5 points

Courbes qui se suivent exactement ou de très près = 10 points pour tous les 500 Hz de coïncidence

Courbes qui se terminent au même endroit = 10 points

Pics ou creux supplémentaires (dans une courbe mais pas dans l'autre) = -5

Pics ou creux qui ne coïncident pas = -5

Impression visuelle de similitude = 0 - 50 points

Chaque paire de courbes a été analysée deux fois avec ce système de points, et les deux résultats se trouvent en haut à droite sur les 7 paires de courbes qui suivent. De ces chiffres de similitude nous pouvons dresser deux listes qui ordonnent les bilingues selon la similitude de leurs deux voix :

- |                     |                     |
|---------------------|---------------------|
| a. Bilingue 1 (130) | a. Bilingue 1 (100) |
| b. Bilingue 5 (55)  | b. Bilingue 2 (87)  |
| c. Bilingue 2 (52)  | c. Bilingue 5 (57)  |
| d. Bilingue 7 (50)  | d. Bilingue 7 (50)  |
| e. Bilingue 3 (45)  | e. Bilingue 3 (33)  |
| f. Bilingue 6 (-40) | f. Bilingue 6 (-40) |
| g. Bilingue 4 (-55) | g. Bilingue 4 (-77) |



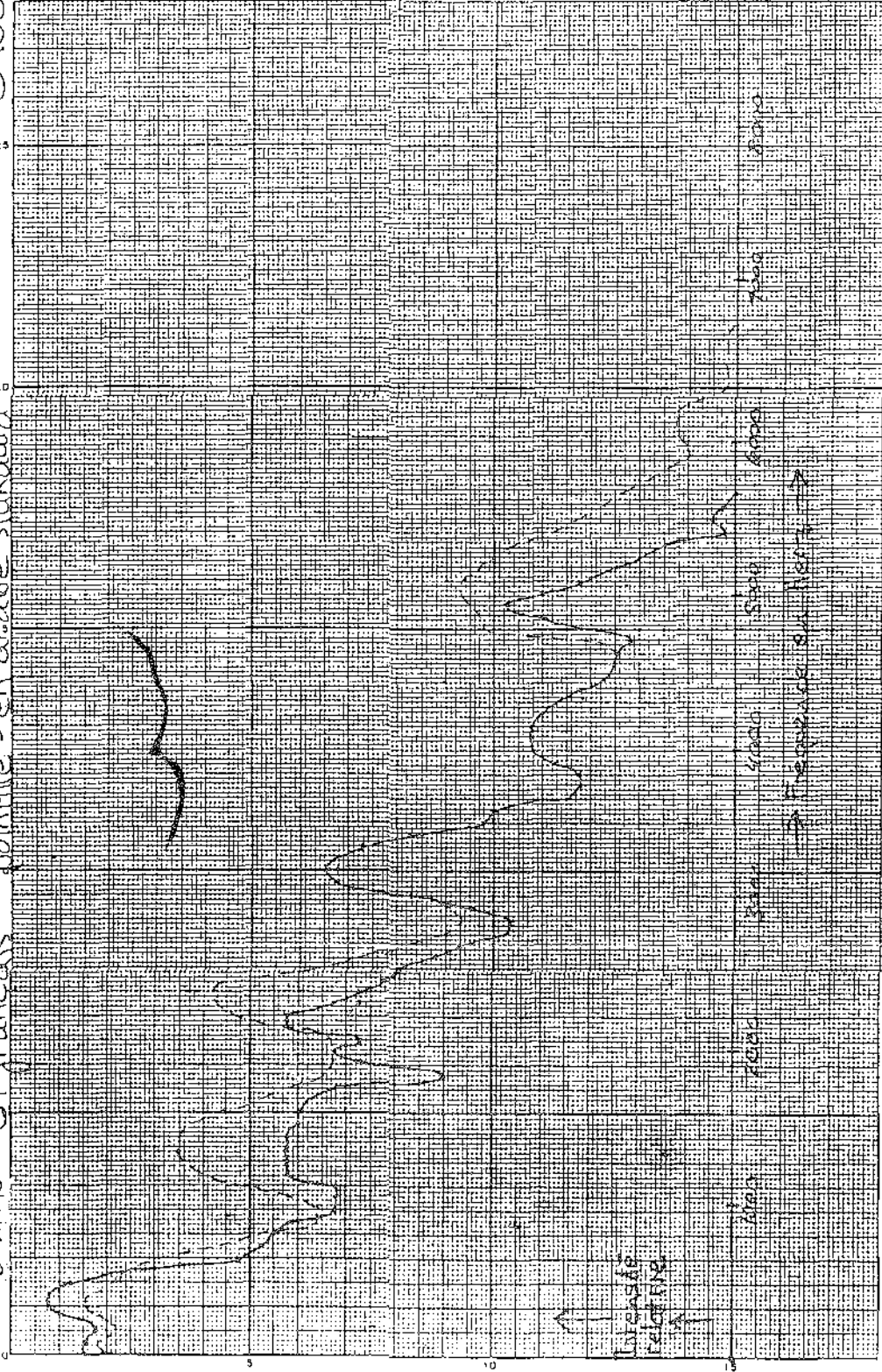
Bilingue n° 1

continu = en français

pointillé = en arabe standard

Points de similitude @ 130

@ 100



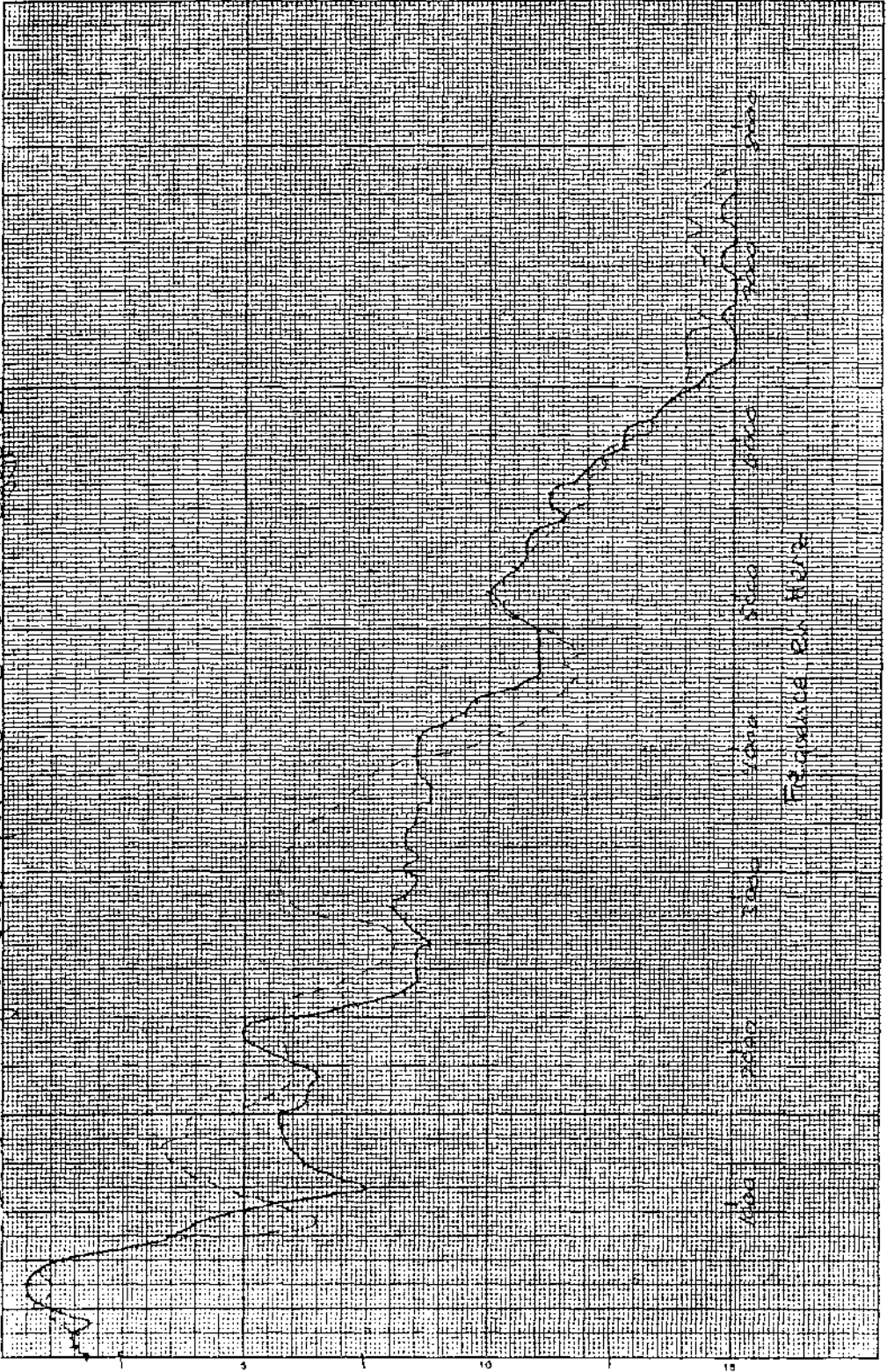
1111

Bilingue n° 2

Points de similitude  $\Phi 52$

$\Phi: 87$

contour = en français pointillé = en arabe égyptien

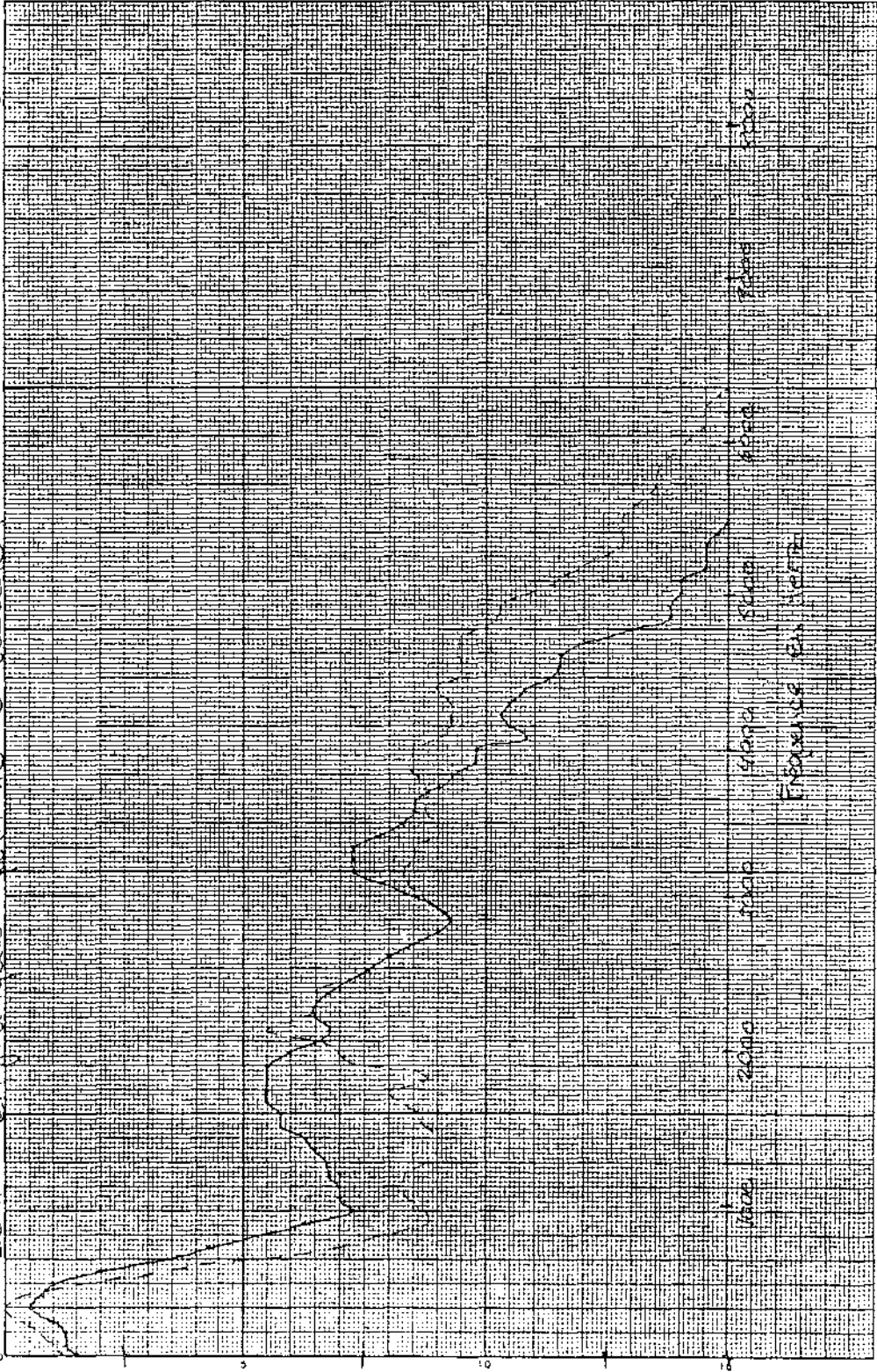


Bilingue n° 3

Ponts de similitude ① 45

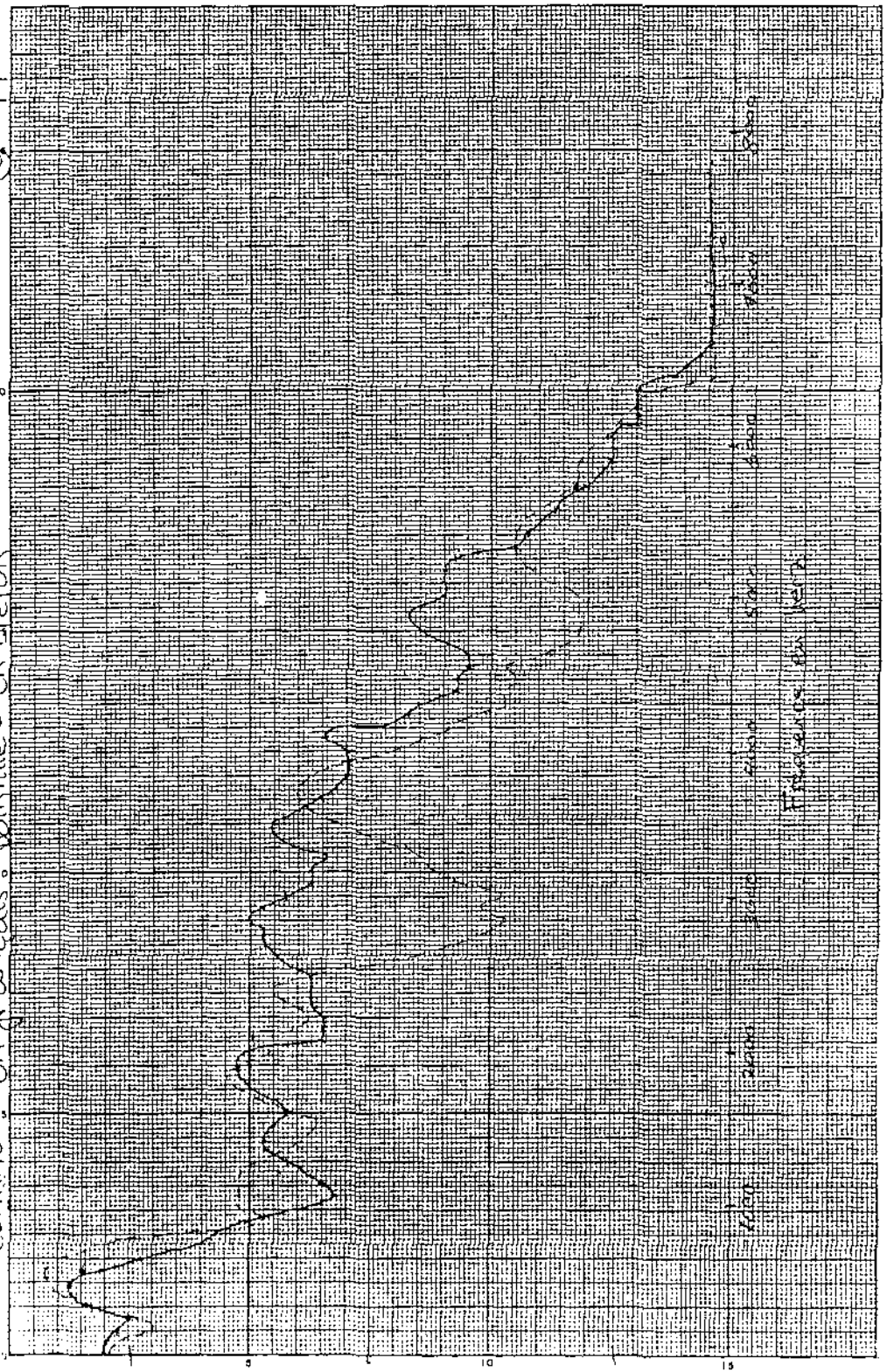
centinu = en français    éantillé = en alsacien

② 33



Bilingue n° 4  
Points de similitude ① -55  
② -77

continu = en français : pointillé = en breton



100  
200  
300  
400

100  
200  
300  
400

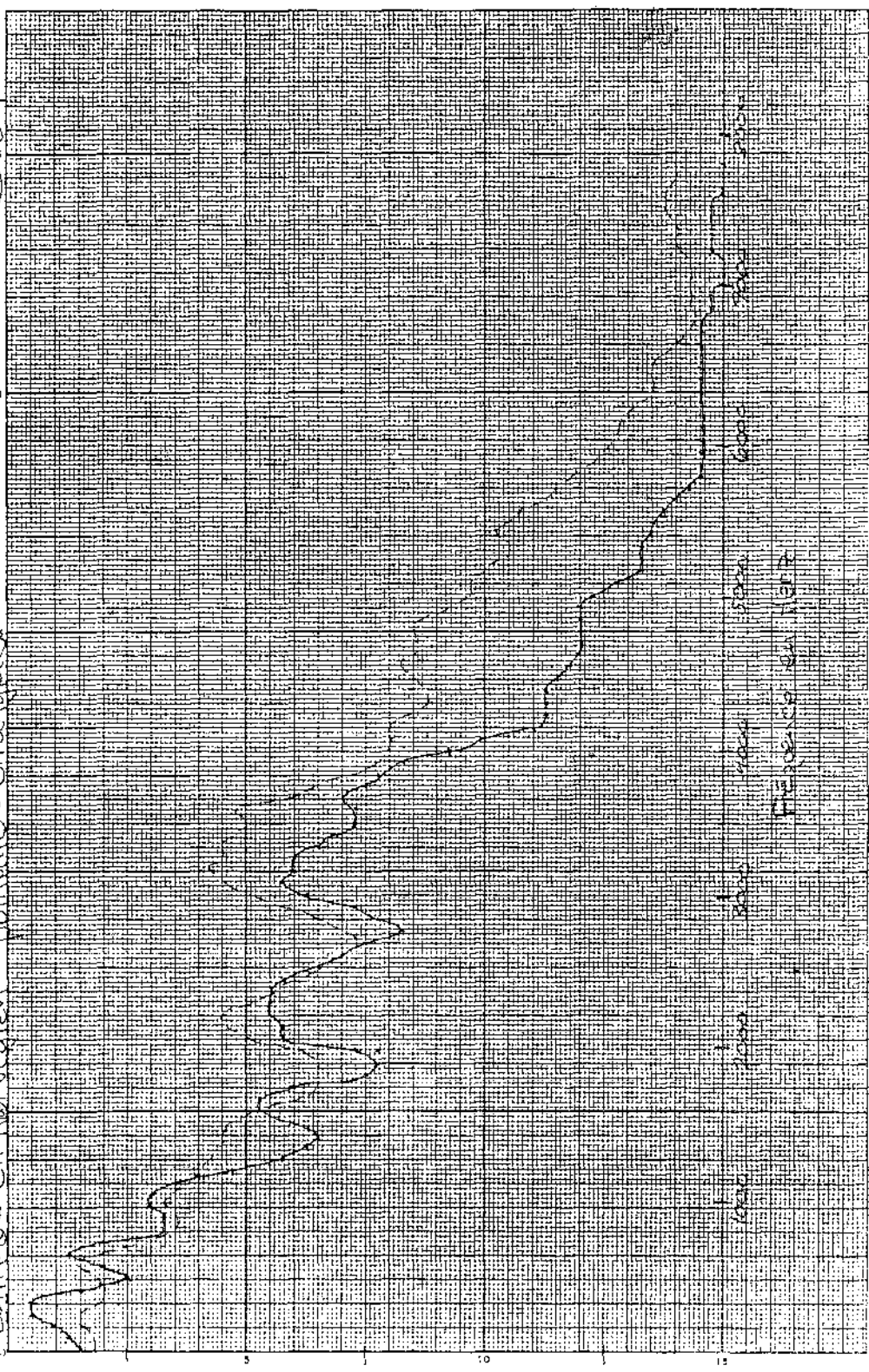


Bilingue n° 5

Points de similitude : ① 55

② 57

continu = en norvégien    pontillé = en anglais



|||||

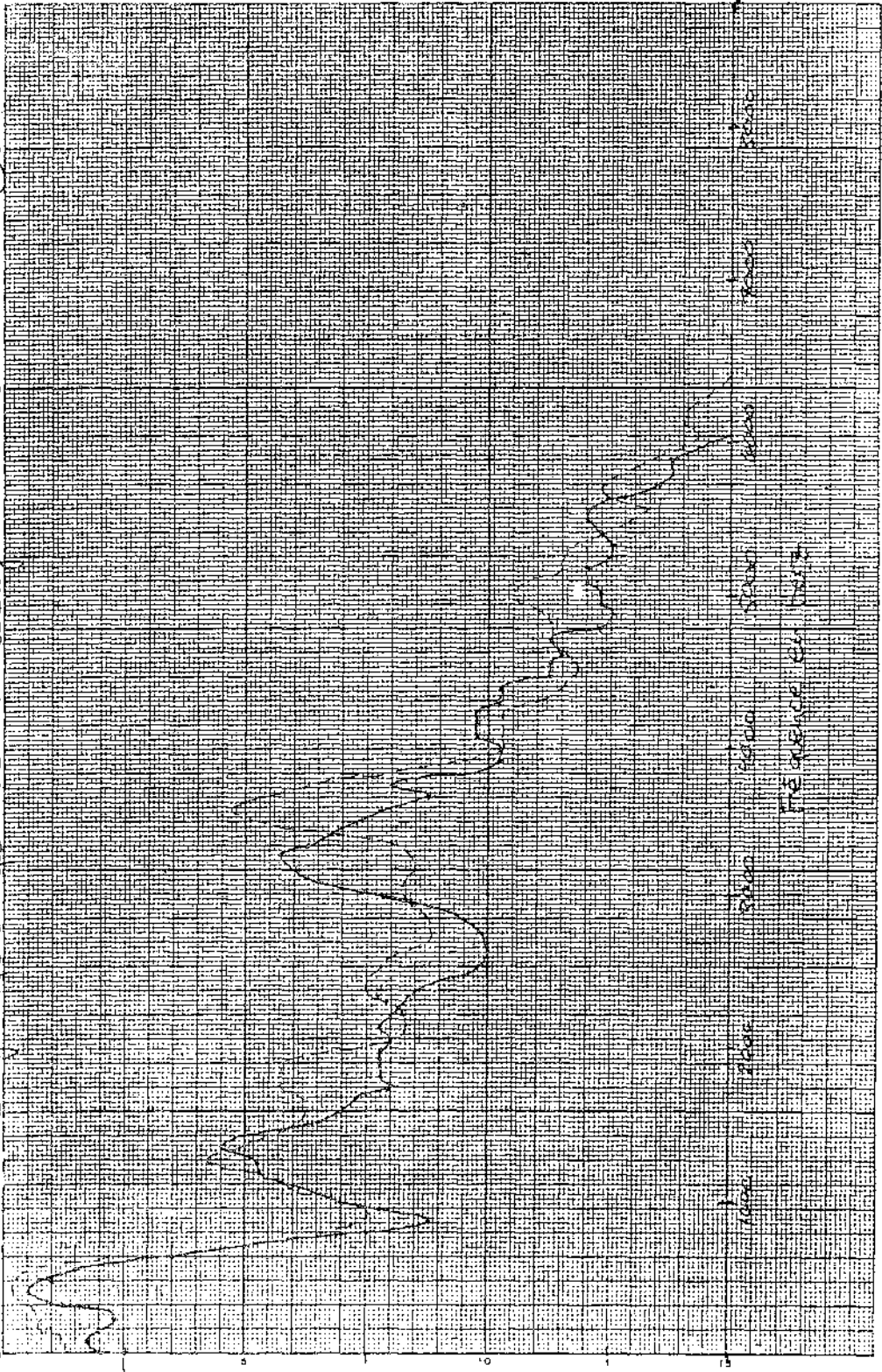
III :

Bilingue n° 6

continu = en français    pontillé = en wolof

Points de similitude ① - 40

② - 40



1000 2000 3000 4000 5000  
0 10 20 30 40 50  
FRANÇAIS EN  
WOLOF

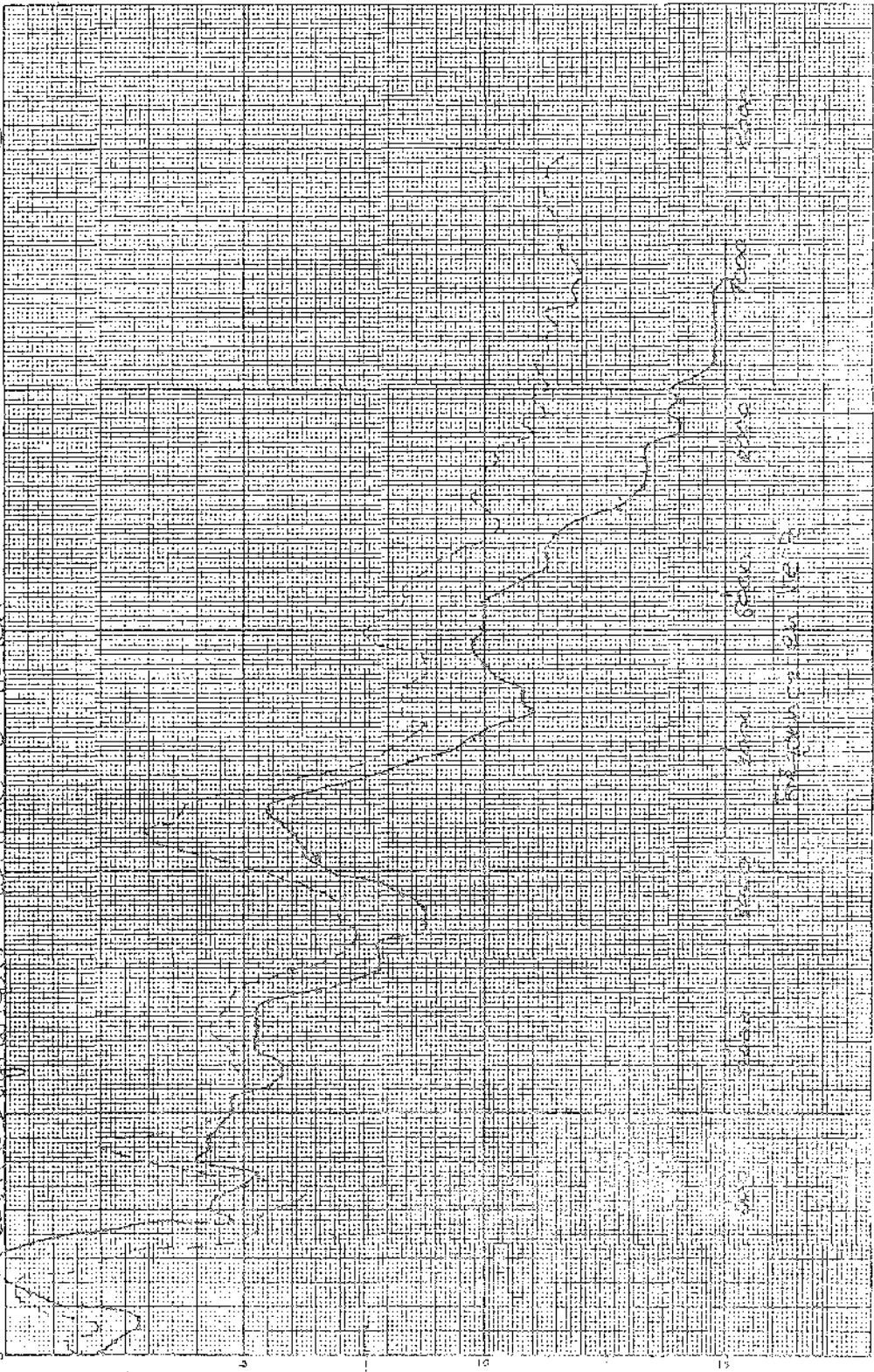
Bilingue n° 7

Points de similitude 1) 50

2) 50

Contour infrancois

contille = en italien



## NOTES

- \* Etude entreprise à l'aide d'un soutien offert par le conseil National des Sciences.
1. TARNOCZY, "Détermination du spectre de la parole avec une méthode nouvelle", *Acustica* 8:392 (1958)
  2. R. SCHILLING, "Stimmuntersuchungen an der Universität Freiburg II", (Leipzig 1929).
  3. R. LUCHSINGER, E. GODFREY, *Voice-Speech-Language*, (Wadsworth Publishing Company 1965).
  4. J.-P. PHELIZON, *Vocabulaire de la Linguistique*, (Editions Roudil 1976).
  6. T. TARNOCZY, "Détermination du spectre de la parole avec une méthode nouvelle", *Acustica* 8:392 (1958)
  7. P. BRICKER, S. PRUZANSKY, "Speaker Recognition", *Contemporary Issues in Experimental Phonetics* 295-326, (Academic Press, N.Y. 1976).
  8. C. BORDONE, R. DUBES, R. PISAMI, G. SACERDOTE, O. TOSI, "Invariances of Talkers Choral Spectra", étude présentée à la 87ème réunion de l'Acoustical Society of America (1974).

## BIBLIOGRAPHIE

- D. ADDINGTON, "The Relationship of Certain Vocal Characteristics with Perceived Speaker Characteristics", Thèse de Doctorat, State University of Iowa (1963).
- B. ATAL, "Automatic Speaker Recognition Based on Pitch Contour", *Journal of the Acoustical Society of America* 52: 1678-97 (1972).
- W. BARTHOLOMEW, "A Physical Definition of 'Good Voice Quality' in the Male Voice", *Journal of the Acoustical Society of America* 6: 25-33 (1934).
- R. BECKER, F. CLARK, F. POZA, R. YOUNG, "A Semiautomatic Speaker Recognition System", Stanford Research Institute Report 1363, (Stanford, California 1972).
- S. BEEMAN, "Semantic Scaling of Voice Quality", Thèse de Doctorat, University of Illinois (1964)
- M. BEZIERS, M. VAN OVERBEKE, *Le Bilinguisme*, (Librairie Universitaire Louvain Cahiers de l'Institut des Langues Vivantes 13).
- R. BOLT, F. COOPER, E. DAVID, P. DENIS, J. PICKETT, K. STEVENS, "Speaker Identification by Speech Spectrograms: Some Further Observations", *Journal of the Acoustical Society of America* 54: 531-4 (1973).
- C. BORDONE, R. DUBEE, R. PISAMI, G. SACERDOTE, O. TOSI, "Invariances of Talkers Choral Spectra", étude présentée à la 87ème réunion de l'Acoustical Society of America (1974).
- P. BRICER, S. PRUZANSKY, "Effects of Stimulus Content and Duration on Talker Identification", *Journal of the Acoustical Society of America* 40: 1331 (1966).
- R. COLEMAN, "Speaker Identification in the Absence of Intersubject Differences in Glottal Source Characteristics", *Journal of the Acoustical Society of America* 53: 1741-43 (1973).
- R. COLEMAN, "Male and Female Voice Quality and its Relationship to Vowel Formant Frequencies", *Journal of Speech and Hearing Research* 14: 565-77 (1973).
- M. DICKENS, G. SAWYER, "An Experimental Comparison of Voice Quality among Mixed Groups of Whites and Negroes", *Southern Speech Journal* 18: 178-85 (1962).
- Dictionnaire de la langue de 19<sup>e</sup> et du 20<sup>e</sup> siècle*, Editions du CNRS (1971).
- C. GRAY, G. KOPP, "Voiceprint Identification", Bell Telephone Laboratories Report, (Bell Laboratories 1944).
- C. HAGEGE, A. HAUDRICOURT, *La Phonologie panchronique*, (Presses Universitaires de France 1978).



- W. HARGREAVES, J. STARKWEATHER, "Recognition of Speaker Identity", *Language and Speech* 6: 63-7 (1963).
- E. HAUGEN, *The Norwegian Language in America: a Study in Bilingual Behavior*, (Indiana University Press 1969).
- H. HOLLEIN, W. MAJEWSKI, "Speaker Identification by Long Term Spectra", *Journal of the Acoustical Society of America* 62: 975 (1977).
- P. LADEFOGED, *Elements of Acoustic Phonetics*, (University of Chicago Press 1962).
- R. LUCHSINGER, E. GODFREY, *Voice - Speech - Language*, (Wadsworth Publishing Company 1965).
- L. KERSTA, "Voiceprint Identification", *Nature* 196: 1253-57 (1962).
- L. KERSTA, "Voiceprint Infallibility", étude présentée à la réunion de l'Acoustical Society of America, Novembre 1962 à Seattle, Washington.
- J.-F. PHELIZON, *Vocabulaire de la linguistique*, (Editions Roudil 1976).
- I. POLLACK, J. PICKETT, W. SUNBY, ?, *Journal of the Acoustical Society of America* 26: 403-6 (1954).
- S. PRUZANSKY, "Pattern Matching Procedure for Automatic Talker Recognition", *Journal of the Acoustical Society of America* 35: 354 (1963).
- S. PRUZANSKY, P. BRICKER, "Speaker Recognition", *Contemporary Issues in Experimental Phonetics* 295-326, (Academia Press, N.Y. 1976).
- R. SCHILLING, "Stimmuntersuchungen an der Universität Freiburg II", (Leipzig 1929).
- R. SCHILLING, "Der Musculus sternothyroideus und seine stimmphysiologische Bedeutung", *Arch. Sprachstimmheilklinik* 1 65 (1937).
- K. STEVENS, C. WILLIAMS, J. CARBONELL, B. WOODS, "Speaker Authentication and Identification: a Comparison of Spectrographic and Auditory Presentations of Speech Material", *Journal of the Acoustical Society of America* 44: 1596 (1968).
- K. STEVENS, A. HOUSE, A. PAUL, "Acoustical Description of Syllabic Nuclei: an Interpretation in Terms of a Dynamic Model of Articulation", *Journal of the Acoustical Society of America* 40: 123
- K. STEVENS, A. HOUSE, "An Acoustical Theory of Vowel Production and Some of its Implications", *Journal of Speech and Hearing Research* 4: 103
- T. TARNOCZY, "Détermination du spectre de la parole avec une méthode nouvelle", *Acustica* 8:392 (1958).
- O. TOSI, R. PISAMI, R. DUBES, A. JAIM, "An Objective Method of Voice Identification", étude présentée à la 7<sup>e</sup> réunion de l'International Association of Voice Identification à New Orleans, Louisiana, (1977).