



La détection de l'intensité de la parole: problèmes et méthodes

Bernard Teston

► To cite this version:

Bernard Teston. La détection de l'intensité de la parole: problèmes et méthodes. Travaux interdisciplinaires du Laboratoire Parole et Langage, 1978, 5, pp.73-137. hal-00285540

HAL Id: hal-00285540

<https://hal.science/hal-00285540>

Submitted on 27 Jun 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TRAVAUX DE L'INSTITUT DE PHONETIQUE D'AIX

Volume 5, 1978. Pp. 73-137

LA DÉTECTION DE L'INTENSITÉ DANS LA PAROLE
PROBLÈMES ET MÉTHODES

BERNARD TESTON
(Université de Provence)

LA DETECTION DE L'INTENSITE DANS LA PAROLE :

PROBLEMES ET METHODES

B. TESTON

I. GENERALITES

La détection d'un phénomène acoustique s'opère au moyen d'appareils de mesure généralement appelés sonomètres. Ce terme englobe toute la chaîne de mesure à savoir : le capteur des variations de pressions acoustiques ou microphone, le détecteur de ces variations de pression, les circuits de pondération et les dispositifs de visualisation et d'exploitation des résultats.

Que mesure-t-on au moyen d'un sonomètre ? Tout simplement un son, c'est-à-dire la manifestation physique d'une variation de pression dans un milieu élastique, variation véhiculant une certaine énergie, dont les grandeurs physiques sont :

- la pression instantanée en un point de l'espace.
- l'intensité, ou flux d'énergie de la variation de pression par unités de surface et de temps.

Dans la pratique, l'aboutissement de la manifestation d'un phénomène acoustique est l'oreille humaine, ce qui impose des restrictions à la définition du son ainsi qu'à sa mesure.

D'une part, nous limitons le son à la bande de fréquences perçues par l'oreille, qui est comprise entre 20 et 20.000 hertz.

D'autre part, la très grande dynamique de la perception auditive associée à un seuil d'audibilité très faible, et la loi de Weber-Fechner, nécessitent une représentation logarithmique qui définit le décibel.

Le décibel n'est pas une unité mais une façon de compter quelque chose, il peut représenter n'importe quoi à condition de se fixer une référence.

Pour l'intensité sonore, cette référence est le seuil d'audibilité de l'oreille moyenne, c'est-à-dire la pression acoustique à laquelle on commence à percevoir une sensation auditive pour un signal de 1000 hertz. Cette pression de référence est de 20 micropascals ou microbars. La mesure physique est ici référencée à partir de données établies empiriquement par les psychophysiologistes.

$$20 \times 10^{-5}$$

Le milieu de perception étant dans le cas général l'air, une source sonore produit donc une pression sonore $p(t)$ qui est une variation de pression dans le temps autour de la pression atmosphérique ambiante. Sa moyenne est donc nulle. Cette variation de pression peut être stationnaire (périodique ou aléatoire) ou transitoire.

On peut lui appliquer une transformation de Fourier et en calculer la valeur quadratique moyenne $\sqrt{p^2(t)}$.

L'intensité sonore est, par définition, le flux d'énergie transporté par cette variation de pression par unité de surface et unité de temps.

Dans le cas d'une onde plane, l'intensité et la pression sont liées par la relation suivante

$$I = \frac{p^2}{\rho c}$$

ρ est la densité de l'air et C la vitesse du son dans l'air. $\rho c = 400$ MKS dans le cas général.

On peut donc définir le niveau de décibel de la pression ou de l'intensité sonore par la relation suivante :

$$L_p = 10 \log \frac{I}{I_0} = 20 \log \frac{P}{P_0}$$

avec $P_0 = 20$ micropascal ou 20×10^{-5} microbar
 $I_0 = 10^{-12}$ Watts par mètre carré.

Les niveaux de pression ou d'intensité acoustique s'expriment en décibels que l'on note dB, généralement suivis d'un indice simple, qualifiant la bande de fréquence dans laquelle la mesure a été effectuée. Par exemple : dB LIN pour linéaire (20-20.000 hertz) ou dBF (de l'anglais flat). Le terme L_p est normalisé, il est issu de l'expression anglaise "sound pressure level".

Pour mesurer un phénomène physique, il est d'abord nécessaire d'en saisir les manifestations au moyen d'un capteur. En acoustique, dans les bandes de fréquences auditives, on utilise un microphone qui transforme les variations de pressions acoustiques en signaux électriques de la manière la plus fidèle possible.

Notre propos n'étant pas d'étudier ces capteurs, nous ne nous étendrons pas davantage sur ce sujet. Les signaux électriques, images de la pression acoustique, sont amplifiés sans distorsion puis détectés, c'est-à-dire transformés en valeurs absolues et intégrés par rapport au temps. Les types de signaux étant très différents, il est préférable de réaliser une détection de valeur efficace. L'intégration est également choisie en fonction de la nature des signaux. Les circuits de pondération sont généralement normalisés et dépendent du type de sonomètre.

La visualisation de la valeur de l'intensité sonore est réalisée soit au moyen d'un indicateur galvanométrique à aiguille sur une échelle logarithmique, soit au moyen d'un enregistreur continu qui permet d'enregistrer les variations d'intensité au cours du temps. Ces appareils ont généralement une dynamique restreinte et leur sensibilité est réglable par paliers afin d'obtenir un fonctionnement satisfaisant des détecteurs. On peut parfois leur adjoindre des filtres pour réaliser des mesures d'intensité sélectives.

II,1) Caractéristiques inhérentes à la mesure de l'intensité des sons du langage :

Pour la mesure des sons de la parole, nous pouvons tout d'abord limiter la bande passante du sonomètre à des valeurs comprises entre 60 et 10.000 hertz. On peut raisonnablement, envisager la limitation de dynamique sur 60 dB Lin (dynamique comprise entre 20 et 80 dB Lin).

Une des principales difficultés réside dans la nature même des signaux de parole. Si leur richesse harmonique se limite à 10 KHz, leur développement temporel est très variable. Le langage est constitué par toutes sortes de signaux ; périodiques, pseudopériodiques, aléatoires, ainsi que des signaux transitoires, lesquels peuvent être parfois mêlés. En fait, les états stables dépassent rarement quelques dizaines de millisecondes et même dans ces conditions, on peut difficilement considérer qu'une période du signal est identique à celle qui la suit ou à celle qui la précède. Ceci, comme nous le verrons plus tard, nécessite des précautions particulières lors de l'élaboration des détecteurs.

La caractéristique la plus importante de l'intensité acoustique du langage est son évolution par rapport au temps. La rapidité des variations de l'intensité au cours de l'élocution rend obligatoire l'enregistrement de ses valeurs instantanées au moyen d'oscillographes enregistreurs rapides pour exploiter les résultats des mesures. D'autre part, si la dyna-

mique de la parole peut être réduite à 60 dB, il n'en demeure pas moins qu'effectuer les mesures sur une telle dynamique sans changement de gamme, nécessite toujours des soins particuliers dans le choix et la réalisation des détecteurs. Pour ces différentes raisons, on a vu apparaître des sonomètres spécialement conçus et réalisés pour des mesures d'intensité de la parole : Boé et Rakotofiringa (1971), Teston (1971) ont décrit des appareils de ce type que les phonéticiens qualifient généralement de détecteurs d'intensité ou d'intensimètres.

II,2) La détection efficace du signal de parole :

Nous appelons détection efficace, la détection de la valeur quadratique moyenne du signal acoustique. L'intérêt particulier que nous portons à la valeur efficace d'un signal ou RMS (Root Mean Square - racine du carré moyen) réside dans le fait que cette valeur est directement liée à l'énergie transportée par le signal par unité de temps. Son expression mathématique est la suivante :

$$X \text{ efficace} = \sqrt{\frac{1}{T} \int_0^T x^2(t) dt}$$

$x(t)$ est le signal électrique image de la pression acoustique fonction du temps.

Il est tout d'abord élevé au carré

$$x(t) \rightarrow x^2(t)$$

Puis intégré sur une période de durée T

$$x(t) \rightarrow x^2(t) \rightarrow \frac{1}{T} \int_0^T x^2(t) dt$$

On réalise, pour finir, la racine du carré moyen du signal.

Cette définition mathématique peut être explicitée plus simplement, en considérant la valeur efficace d'un signal de forme quelconque comme la valeur d'un signal continu qui produit la même énergie pendant le même temps. Cette notion est très importante, car elle permet de comparer avec sûreté des signaux divers en comparant leur énergie.

Lorsqu'on étudie des signaux dans le domaine temporel, et que l'on s'intéresse particulièrement à leur amplitude, on peut définir plusieurs références de comparaison :

- L'amplitude crête à crête,
- L'amplitude crête,

- L'amplitude moyenne,
- L'amplitude efficace.

Si l'on mesure un son pur sinusoïdal, on peut calculer les différentes amplitudes les unes par rapport aux autres, car elles ne diffèrent que par des constantes multiplicatives. (Figure 1)

Dans ces conditions, un détecteur de valeur efficace pour son pur se réduit à un détecteur de valeur absolue moyenne équipé d'un rapport multiplication de I, II. Malheureusement, ce cas est assez artificiel et très rare !

Pendant longtemps, on s'est contenté de réaliser des mesures de valeur moyennes en redressant les deux alternances des signaux par rapport au 0 électrique, et en leur appliquant une constante de temps d'intégration. Certains de ces détecteurs sont appelés détecteur quasi RMS. Pour les sonomètres de mesures acoustiques, on emploie surtout des circuits d'élévation au carré constitués par des quadrateurs à diode (Wahrman 1958) qui n'ont cessé d'être améliorés (Hansen 1972 - Brie 1971), ces détecteurs donnent de bons résultats mais manquent de dynamique. Ils sont particulièrement bien adaptés aux sonomètres dont la dynamique est divisée en plusieurs gammes, ce qui leur permet de fonctionner dans de bonnes conditions. Pour des dynamiques supérieures à 40 dB Lin, on leur adjoint généralement des circuits de commutation automatique ou de contre réaction qui leur font perdre certaines de leurs qualités (facteur de crête). Après l'élévation au carré du signal et l'intégration, la racine carrée ainsi que la conversion logarithmique sont généralement réalisées par la dynamique du galvanomètre de lecture. Parfois même on s'abstient de réaliser la racine carrée, ce qui donne à la dynamique de sortie logarithmique une valeur double ($\log X^2 = 2 \log X$). L'évolution de la technologie des composants actifs a permis la construction, depuis quelques années, de modules baptisés convertisseurs efficace-continu (RMS - DC) qui mesurent la température d'un élément thermique chauffé par le signal. Comme convertisseur de valeur efficace, on ne fait théoriquement pas mieux. Mais ces dispositifs, sont caractérisés par une faible aptitude à répondre à de rapides variations d'amplitude, malgré une bande passante parfois élevée et un facteur de crête important. D'autres types de convertisseurs simulent au moyen du calcul analogique l'équation de la valeur efficace (Figure 2).

Le calcul analogique est opéré au moyen de circuits logarithmiques, antilogarithmiques et d'amplificateurs opérationnels, réalisant des gains fixes pour simuler les multiplieurs et diviseurs nécessaires au calcul de la valeur efficace du signal d'entrée. Ces dispositifs ont une bande passante importante, une dynamique de l'ordre de 60 dB, avec une précision inférieure à 1 % et un facteur de crête qui peut s'élever jusqu'à 10 dans le meilleur des cas. Ces détecteurs sont parfois appelés LMS (Logarithmic Mean Square).

Il nous faut définir maintenant le facteur de crête d'un détecteur de valeur efficace. C'est une caractéristique primordiale pour ces dispositifs. Le facteur de crête n'est pas une propriété du détecteur mais une possibilité, il caractérise, en fait, le signal dont on veut connaître la valeur efficace. Mathématiquement le facteur de crête d'un signal est le rapport de sa valeur crête sur sa valeur efficace.

$$\text{Facteur de crête} = \frac{A_{\text{crête}}}{A_{\text{efficace}}}$$

Pour un échelon de tension il est égal à 1 ($A_{\text{crête}} = A_{\text{efficace}}$).

Pour une sinusoïde il est de 1,414 ($\sqrt{2}$).

Nous exposons, dans le tableau qui suit, quelques facteurs de crête mesurés sur des voyelles.

Voyelle	différence crête/RMS	facteur de crête
a	8 dB	2,5
i	6 dB	2
o	5 dB	1,77
y	7 dB	2,23
u	6 dB	2

On en déduit que le facteur de crête se maintient, pour des voyelles dans des limites raisonnables (entre 1,5 et 3) mais qu'il représente tout de même des variations du simple au double.

Pour les consonnes constrictives, le facteur de crête est déjà plus élevé

ch [S]	12 dB	4
s [s]	14 dB	5

Pour les bruits d'explosion, sa valeur peut être encore plus importante (Nous avons relevé une explosion de (K) de 18 dB de différence, donc d'un facteur de crête de 8, mais cette valeur est exceptionnelle). Pour les bruits d'explosion des consonnes occlusives il se situe généralement plus bas, de l'ordre de 6 à 7.

Dans ce cas, se greffe un problème supplémentaire qui est particulier à la détection efficace de ce type de signaux transitoires. En effet, l'erreur d'évaluation de la mesure efficace de ces signaux est inversement proportionnelle au facteur de crête du détecteur et au rapport durée de l'impulsion / durée d'intégration.

Pour conclure, nous pouvons considérer qu'un détecteur de valeur

efficace capable de mesurer des signaux d'un facteur de crête de 7, peut être considéré comme satisfaisant pour la mesure des signaux acoustiques de la parole.

II,3) La constante de temps d'intégration :

La définition mathématique de la détection efficace fait intervenir une intégration du signal par rapport au temps. Pour son calcul, on doit tenir compte du développement du signal dans le temps. Pourquoi ? Parce qu'il est nécessaire, pour décrire un phénomène, de l'observer pendant un certain temps. La description la plus précise d'un signal qui se développe dans le temps consiste à l'enregistrer complètement. C'est l'intégration moyenne sur un temps long, où le temps d'intégration T est choisi égal au temps d'observation du signal. Cependant, cela n'est pas toujours nécessaire, tout le problème étant de savoir ce que l'on veut effectivement étudier dans le signal. On peut donc utiliser, l'intégration moyenne par palier, où le signal est intégré sur un temps T de manière continue. Chaque intégration commence à la fin du temps T précédent. Le résultat de l'intégration est donné à la fin de chaque temps T . Il est possible encore d'utiliser l'intégration continue qui est réalisée sur les dernières T secondes du signal de manière continue c'est-à-dire que l'intégration est réalisée au moyen d'une fenêtre temporelle que l'on déplace au cours de l'évolution du signal dans le temps.

La valeur de cette intégration varie de manière continue, contrairement à l'intégration par palier, où le résultat d'une période d'intégration T est indiqué et stable, jusqu'à ce que soit terminée l'intégration en cours. Enfin, on peut utiliser l'intégration moyenne pondérée. C'est le type d'intégration le plus utilisé. Elle est réalisée de manière fort simple, au moyen d'un circuit analogique intégrateur RC.

Outre sa simplicité de réalisation, l'intégration moyenne pondérée, encore baptisée exponentielle, a des fronts de montée et de descente qui simulent au mieux les réponses de l'oreille aux attaques et aux fins de salves de signaux, respectivement : - temps d'établissement et durée de réverbération (Von Békésy 1960).

Wahrman et Broch (1975) ont comparé les résultats de l'intégration continue et de l'intégration moyenne pondérée pour les différentes classes de signaux, c'est-à-dire : stationnaires périodiques ou aléatoires, transitoires ou impulsionnels.

Si la méthode d'intégration continue est mathématiquement idéale, on peut considérer que l'intégration moyenne pondérée lui est équivalente si :

$T = 2 RC$ pour les signaux stationnaires
 $T = RC$ pour les signaux impulsionnels, T étant la période d'observation du signal.

Comment la choisir ? En fait, tout le problème se situe dans ce choix. Dans le cas d'un signal périodique stationnaire, il est nécessaire, pour en avoir une bonne description, de l'observer pendant au moins une période. On en déduit T . Pour un signal sinusoïdal qui est symétrique, on peut même se contenter d'une demi-période, mais ceci est un cas particulier. Pour les bruits aléatoires, le choix de T se fait à la lumière de la largeur de bande du bruit et de sa composante fréquentielle la plus basse.

$$T = \frac{1}{\pi f_0} = 2 RC$$

f_0 étant la fréquence de coupure à 3dB du réseau RC de l'intégrateur. Dans le cas des signaux transitoires et impulsionnels, T doit être choisi avec beaucoup plus d'attention. Si l'on considère le rapport de la durée de l'impulsion sur la période d'intégration $\frac{t_i}{T}$, on remarque que le minimum d'erreur par rapport à l'intégration idéale, se situe pour une valeur de $\frac{t_i}{T} = 0,1$ c'est-à-dire pour une constante de temps $T = RC$ égale à 10 fois la durée de l'impulsion. Ceci pour un détecteur de valeur efficace capable de détecter des signaux d'un facteur de crête de 5.

Avec $0,03 T < t_i < 0,3 T$ l'erreur est de l'ordre de 0,5 dB. Cette observation est importante, car il n'est pas envisageable de faire varier T pour chaque type de signaux à mesurer et il faut tenir compte de ces possibilités d'erreur pour choisir une constante de temps qui constitue un bon compromis. Quelle est donc la valeur que nous devons donner à la constante de temps d'intégration d'un détecteur de valeur efficace pour mesurer les signaux de parole ? Les critères que nous venons de développer précédemment font apparaître :

- que l'intégration moyenne pondérée ou exponentielle est satisfaisante pour tous les types de signaux que nous avons à mesurer dans la parole.
- que le choix de la constante de temps doit tenir compte de la détection des transitoires, il faut pour cela que cette constante ne soit ni trop grande ni trop petite par rapport à la durée de ces derniers.
- que la constante de temps ne doit pas être trop faible pour la mesure des sons voisés afin d'éviter des oscillations résiduelles (ripple) trop importantes, qui sont génératrices d'erreurs.
- que, en revanche, la constante de temps ne doit pas être trop importante, pour ne pas engendrer des erreurs de mesure compte tenu des fortes et rapides variations d'amplitude du signal pendant l'élocution.

Toutes ces précisions étant apportées, il nous faut rechercher une valeur de compromis idéale.

II,4) Valeur de la constante de temps d'intégration :

a) Cas des segments voisés :

Nous groupons, sous cette rubrique, les voyelles et les consonnes voisées. Il est établi que le maximum d'énergie se trouve dans le bas du spectre pour les voyelles. C'est donc par rapport au fondamental, qui par définition est la fréquence la plus basse que nous devons choisir la constante de temps d'intégration (Fig. 3).

Si l'on se fonde sur le critère d'équivalence défini précédemment pour les signaux périodiques stationnaires $T = 2 RC$ et, si l'on considère la fréquence de 100 hertz comme fréquence la plus basse, nous aurons :

$$T = 10 \text{ ms} \qquad RC = \frac{T}{2} = 5 \text{ ms}$$

Ceci pour un son pur et symétrique ou une salve de signaux périodiques de durée "infinie". Le signal de parole, pouvant être fortement dissymétrique, et ne répondant pas à ces conditions, une constante de temps $RC = 10 \text{ ms}$ nous semble mieux adaptée (Fig. 6).

b) Cas des consonnes constrictives :

Ce type de signaux est essentiellement constitué par du bruit, avec une répartition spectrale marquée par des pôles d'énergie et un faible niveau dans les basses fréquences. Le critère d'équivalence est le même que pour les sons voisés : $T = 2 RC$.

On peut donc considérer pour ce type de signaux une constante de temps $RC = 10 \text{ ms}$ comme satisfaisante (Fig. 7).

c) Cas des consonnes occlusives :

Ces consonnes se manifestent essentiellement par des bruits d'explosion qui sont des transitoires. Nous avons vu que ce type de signaux est le plus difficile à détecter et intégrer correctement. En fait tout le problème réside dans la durée de ces transitoires (Kim et Mc Neilage en 1972). Sans avoir fait des mesures systématiques, nous avons observé les valeurs suivantes :

K : 6 à 8 ms selon le contexte

p : 4 à 7 ms

t : 3 à 4 ms

Nous voyons donc que pour satisfaire la relation $0,03 T = t_i$ $0,3 T$, $T = 3 t_i = 25$ ms dans le pire des cas ; donc $RC = 25$ ms.

Cependant, si l'on observe bien le développement temporel de ces signaux, on s'aperçoit que lorsque la durée est supérieure à 5 ms ils sont constitués de plusieurs impulsions successives.

Avec une constante de temps de 10 ms, l'erreur ne doit pas être supérieure à 1 dB. La valeur de $RC = 10$ ms semble donc être un bon compromis (Fig. 4 B).

d) Problèmes liés à la variation temporelle de l'amplitude des signaux du langage.

Nous avons déjà mentionné rapidement les grandes variations de dynamique des signaux de parole au cours de l'émission du message verbal. Ces variations peuvent être également très rapides. On peut assimiler les signaux de parole à des salves tonales et à des bruits, entrecoupés de silences et de transitoires. Pour montrer l'influence de la constante de temps sur une telle suite de signaux, on peut simuler ces phénomènes par des trains de signaux périodiques ou "tone burst". On voit immédiatement, l'effet sur ces signaux, d'une intégration trop forte (Fig. 4 A). Si la mesure de l'amplitude des signaux est bonne, on peut ne pas s'apercevoir qu'à certains instants ils sont interrompus. Ceci est dû au fait que l'intégration par moyenne pondérée ou exponentielle met un certain temps pour revenir à zéro lorsque le signal disparaît, car le front de descente de l'intégrateur est une exponentielle. On peut considérer ce phénomène comme une réverbération. Comme on associe toujours la détection efficace à une conversion logarithmique, le front de descente est une droite dont on peut évaluer le temps de chute en décibels par seconde.

On applique pour cela la formule suivante :

$$\frac{I}{RC} = \frac{\text{dB seconde}}{8,7}$$

(fig. 5).

Les variations d'amplitude des signaux du langage proviennent des différences intrinsèques entre les segments contigus et des facteurs prosodiques tels que l'accent, l'expressivité, le contraste (Lehiste 1970 et Klatt 1976). Tous ces paramètres varient dans des proportions importantes d'une langue à l'autre, comme au sein d'une même langue. Les effets de l'intégration dans une séquence équivalente à une phrase se distinguent particulièrement dans les exemples que nous donnons dans les figures 8, 9 et 10.

On en déduit qu'une constante de temps de 10 ms représente ici encore un bon compromis.

En fait, la valeur d'une constante de temps comprise entre 10 et 20 ms, est utilisée par les phonéticiens depuis un certain temps et d'une manière empirique, car c'est autour de ces valeurs que l'on obtient une représentation de la courbe d'intensité suffisamment précise et qui n'entraîne pas un lissage trop important des variations d'intensité.

III,1) Les pondérations subjectives :

Il est bien connu que l'oreille ne perçoit pas toutes les fréquences acoustiques avec la même intensité. Cette perception est fonction de l'énergie du signal. Ces variations perceptives sont représentées sur des graphiques baptisés courbes isosoniques ; elles ont été établies pour la première fois par Fletcher et Munson (1933) au moyen de tests psycho-acoustiques portant sur des sons purs. Ces courbes sont normalisées sous l'appellation NF S 30-003.

Les courbes isosoniques représentent les niveaux sonores de sons purs qui produisent sur un sujet normalement bien entendant la même sensation d'intensité sonore. Cette sensation est appelée sonie, en anglais "loudness". La référence de ces courbes isosoniques est donnée à la fréquence de 1000 hertz, fréquence à laquelle l'oreille suit bien la loi logarithmique. Pour définir le niveau d'isotonie d'un phénomène sonore, on demande à des sujets de comparer les niveaux d'intensité de ce phénomène avec un son de référence de 1000 hertz.

Lorsque les sujets ont égalisé les deux sensations auditives, la valeur de la pression acoustique à 1000 hertz donne le niveau d'isotonie en PHONES. Le PHONE est une unité sans dimension, utilisée pour évaluer le niveau d'isotonie d'un son. Le niveau d'isotonie est donc dit de X PHONES, lorsque la sonie ou sensation auditive d'un son est jugée équivalente par un sujet normal, à celle d'un son pur de 1000 hertz se propageant en ondes planes face au sujet, et dont le niveau de pression acoustique est de X dB au-dessus du seuil d'audibilité (20 microbars).

L'échelle des phones est donc logarithmique comme pour les décibels. On peut utiliser également une échelle linéaire qui définit le SONE. Par définition : 1 SONE = 40 PHONES. On double les SONES tous les 10 PHONES à partir de 40, en suivant les puissances de 2 :

40 PHONES	:	1 sone
50 "	:	2 "
60 "	:	4 "
70 "	:	8 "
80 "	:	16 "

La représentation en SONES est surtout utilisée comme artifice de calcul.

Dans la pratique, pour mesurer la sonie d'un son, on utilise des appareils de mesure objectifs appelés sonomètres. Ils sont sensés nous donner une information qui doit être la plus proche de la sensation auditive.

Comme la sensation auditive est liée à la fréquence mais qu'elle n'est pas linéaire par rapport à la pression acoustique, on définit des pondérations subjectives dérivées des courbes isosoniques et normalisées. Les courbes les plus utilisées sont les références A, B, C, qui correspondent respectivement aux niveaux de 40, 70 et 100 Phones.

Si l'on étudie en détail les courbes isosoniques, on remarque qu'elles ne correspondent pas tout à fait aux courbes de pondération normalisées (Fig. 11 et 12). La courbe A, par exemple, ne tient pas assez compte de l'accroissement de sensibilité de l'oreille entre 2000 et 6000 hertz. Nous avons donc réalisé une courbe de pondération qui augmente de 4 dB à 4000 hertz la courbe A normalisée (Fig. 13), mais les résultats obtenus au moyen de ces deux courbes sont très voisins.

III,2) Pondération en fonction du type de champ acoustique :

La sensation auditive dépend également de la direction du son. On se contente, en général, de faire des mesures en ondes planes ou champ frontal. Pour des mesures en champs diffus, on a essayé des pondérations globales qui consistent à augmenter les hautes fréquences en simulant approximativement les caractéristiques de ce type de champ acoustique, mais cette méthode ne donne pas de résultats satisfaisants. Voyons ce qu'il en est exactement pour la mesure des signaux de parole.

Considérons tout d'abord la source du rayonnement acoustique du langage c'est-à-dire la bouche. Flanagan (1960) a étudié la radiation sonore de la bouche et en a déduit les résultats suivants :

- la distribution de la pression acoustique est relativement uniforme dans l'axe de la bouche. On peut considérer cette dernière comme une source ponctuelle générant un champ sphérique. (en anglais "far field") avec comme limites de droite à gauche +40,0, -40 et de haut en bas +20,0, -20 degrés.

- la variation de la pression acoustique en fonction de la distance est identique à celle de toute source ponctuelle créant un champ sphérique. Sa perte est de 6 dB chaque fois que l'on double la distance de la source au capteur de mesure.

- La pression sonore à une distance de 35 centimètres dans l'axe de la bouche est caractérisée par une augmentation de son amplitude de + 7dB par octave, à partir de 250 hertz.

Nous pouvons déduire de ces constatations que, à quelques dizaines de centimètres dans l'axe de la bouche, un microphone de petite dimension (2,5 centimètres de diamètre) capte approximativement une onde plane.

Dans les conditions d'enregistrement normales, on n'a donc pas à opérer des pondérations particulières. La prise de son se fait à 50 centimètres de la bouche face au sujet et en champ libre dans le cas idéal.

Une distance plus proche augmente les bruits de souffle et restreint l'hypothèse d'onde plane en diminuant le rayon de la sphère du champ acoustique émis par la bouche. La radiation aux lèvres augmentant le niveau des fréquences aigües, il n'est pas utile de recourir à une pondération particulière si ce n'est pour connaître, à partir de la mesure, la pression acoustique aux lèvres du locuteur.

III,3) Pondérations en fonction de la structure acoustique :

Plus délicates sont ces pondérations à apporter au signal pour l'évaluation de la sonie du langage. Nous avons décrit précédemment, une mesure au moyen des courbes normalisées. Les figures 14, 15, 16, 17 et 18 nous montrent des résultats très différents pour la mesure de l'intensité. Peut-on considérer ces pondérations comme valables ? Non, car les tests de perception montrent qu'un son complexe n'est pas perçu à la sonie mesurée au moyen de ces courbes de pondération. On mesure généralement une valeur plus faible de la sonie. On a essayé d'améliorer les mesures subjectives, en faisant varier les courbes de pondération en fonction de l'intensité acoustique détectée par les sonomètres, mais elle ne donne pas de meilleurs résultats. En fait, en pondérant l'intensité des signaux du langage au moyen des courbes isosoniques on ne fait pas intervenir tous les paramètres régissant les mécanismes de la perception auditive. Les courbes isosoniques sont établies avec des sons purs ; elles ne sont donc valables que pour la perception de sons purs ou à bande de fréquence étroite. Pour effectuer une mesure de la sonie plus réaliste, il est nécessaire de faire une analyse spectrale du signal, car la sonie des sons complexes est fonction de leurs composantes spectrales. Elle est également fonction de l'effet de masque, phénomène qui est mis en évidence par le fait que la sonie de deux bandes de fréquence entendues simultanément est inférieure à la somme des sonies des deux bandes entendues séparément. Les principales méthodes de calcul de la sonie connues actuellement sont celles définies par Stevens (1961) et Zwicker (1960).

Elles utilisent l'analyse en fréquence par 1/3 d'octave ou par bandes critiques. Les deux méthodes donnent des résultats relativement proches et sont maintenant normalisées pour le calcul de la sonie des bruits continus. La mesure de l'intensité subjective ne peut donc être réalisée au moyen d'un intensimètre qu'en apportant à une mesure objective un facteur de pondération F qui tient compte de tous les phénomènes énumérés précédemment. On peut donc, à partir de courbes d'intensité objectives, calculer les valeurs subjectives manuellement au moyen des différentes méthodes de calcul de la sonie. Pour la sonie des voyelles, Rossi (1971) propose une méthode dérivée de celle de Stevens (1961). Pour la sonie des consonnes, nous ne connaissons pas une méthode particulière de calcul mais l'on peut considérer la méthode de Zwicker (1960) comme la plus fondée pour l'évaluation de la sonie des bruits.

III,4) Pondérations en fonction de la durée des phénomènes acoustiques.

L'intensité effectivement perçue d'un phénomène acoustique varie en fonction de sa durée. On peut définir plusieurs constantes de temps ou durées critiques dans les mécanismes de la perception auditive. Premièrement, la valeur de quelques dizaines de millisecondes environ, qui est la durée moyenne du temps de reconnaissance d'un son. Celui-ci varie en fonction de la durée du front d'établissement, de sa forme et de sa fréquence. Cette durée est à rapprocher de la valeur de 50 millisecondes de la période d'un phénomène acoustique au-dessous de laquelle il est perçu comme un son continu et au-dessus de laquelle il est perçu comme une suite d'impulsions ou de battements (20 hertz). Cette valeur a été parfois prise comme constante de temps d'intégration des détecteurs pour la mesure de l'intensité de la parole ; nous avons vu précédemment, qu'elle est trop élevée pour cette application. La seconde durée critique est de l'ordre de 200 millisecondes durée nécessaire pour que l'oreille perçoive toute l'information contenue dans un stimulus acoustique (dans la parole, rares sont les segments acoustiques stables supérieurs ou égaux à cette durée). La troisième constante de temps critique se situe autour de la valeur de 20 millisecondes durée en dessous de laquelle l'oreille ne perçoit plus les modulations d'amplitude d'un signal ("roughness" en anglais, que l'on peut traduire par rugosité). La quatrième et dernière durée critique a une valeur de 2 millisecondes. Elle constitue la durée minimale de discrimination entre deux structures acoustiques perçues comme différentes. Au-dessous de cette valeur les deux structures acoustiques sont perçues comme un stimulus unique. Cette valeur correspond également à la durée minimale d'une structure acoustique qui peut influencer la perception d'un stimulus dans lequel elle est contenue (Zwicker, 1972 - Green, 1972).

Ces quatre durées critiques influencent de manière variable et à des niveaux très différents, la perception effective du signal de parole, compte-tenu du développement temporel des sons du langage et des mécanismes d'intégration de l'audition en fonction de la durée des segments acoustiques (Zwislocki, 1969).

Il est possible d'effectuer certaines pondérations, après avoir mesuré la durée des différents segments acoustiques ainsi que leur pente de montée ou de descente, et d'en définir une correction (ROSSI, 1976), mais à condition que la durée du segment soit supérieure à 200 millisecondes. Réaliser une pondération automatique avec ce type de paramètre semble plus compliqué qu'avec les paramètres acoustiques.

IV,1) Réalisation de l'intensimètre :

Nous avons réalisé un intensimètre en tenant compte des considérations précédemment développées, qui répond aux caractéristiques suivantes :

- Dynamique de 60 dB,
- Détection de valeur efficace d'un facteur de crête de 8 au minimum,
- Constante de temps d'intégration exponentielle de 5-10 et 20 millisecondes,
- Pondération selon les courbes A-B et C,
- Calibration à 1000 hertz par bonds de 1 dB.

Nous l'avons donc réalisé suivant le schéma de la figure 19.

IV,2) Préamplificateur :

Pour faire fonctionner le détecteur de valeur efficace au maximum de sa dynamique, en utilisant comme source tout type de magnétophone, l'intensimètre est équipé d'un préamplificateur de 10 K ohms et d'impédance d'entrée et d'un gain total de 40 dB. Il est constitué par deux amplificateurs opérationnels LF 356 N de 20 dB de gain chacun encadrant un potentiomètre de 10 K ohms logarithmique qui permet de régler le niveau du signal d'entrée.

IV,3) Oscillateur de calibration :

Cet oscillateur délivre un signal de 1000 hertz pour opérer la calibration de l'enregistreur galvanométrique avant chaque manipulation. De par sa fonction cet oscillateur doit être stable en amplitude. Pour cela, nous utilisons un circuit LM 170 qui contient un contrôle automatique de gain permettant de stabiliser son signal de sortie à une valeur meilleure que 0,5 % (SMATHERS - 1971). L'oscillateur est suivi d'un amplificateur qui permet d'obtenir un signal de sortie d'une valeur de 20 volts crête à crête. L'oscillateur est associé à un atténuateur constitué par deux cellules en échelles, l'une de 6 pas de 10 dB, l'autre de 9 pas de 1 dB.

IV,4) Filtre de Pondération :

Bien que nous ne préconisons pas leur emploi, nous avons équipé l'intensimètre des filtres de pondération A, B et C car ils sont normalisés et présents sur la majorité des sonomètres. Ils sont réalisés au moyen de simples réseaux passe-haut et passe-bas RC adaptés en impédance au moyen d'amplificateurs opérationnels LM 301 A.

IV,5) Détecteur de valeur efficace :

Notre choix s'est porté, pour l'organe essentiel de l'intensimètre, sur un module hybride LH0091 CD de NATIONAL SEMICONDUCTOR.

Ce détecteur de valeur efficace permet dans une gamme de tension

d'entrée comprise entre 50 mV et 7 volts RMS une erreur de $1 \text{ mV} \pm 0,2 \%$. La bande passante pour ce taux d'erreur est dans le pire des cas (avec $V_e = 10$ volts de 15 K hertz (à - 3 dB sa bande passante est de 100 k hertz) (Fig. 21). Il peut accepter des signaux d'un facteur de crête de 8 avec une erreur de mesure inférieure à 0,2 %. Il correspond bien aux caractéristiques que nous avons définies pour la mesure de l'intensité des signaux de parole. La constante de temps est réalisée au moyen d'un simple condensateur que l'on ajoute au circuit. Elle est effectuée par moyenne pondérée (exponentielle et a pour valeur $T = 1/2 RC$ (avec $R = 10 \text{ K ohms}$) car ce détecteur de valeur efficace est un dispositif à contre-réaction. Nous avons utilisé ce composant dans sa configuration dite à fort facteur de crête dans la configuration préconisée par le constructeur ; il nous donne ainsi toute satisfaction.

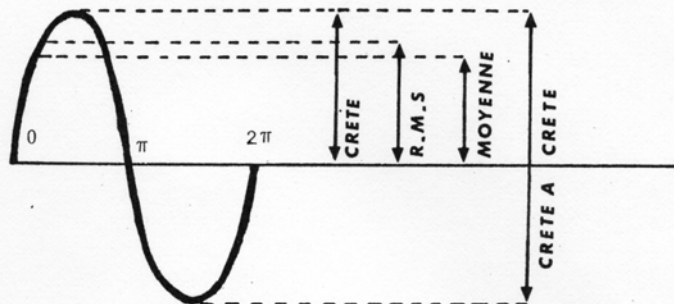
IV,6) Amplificateur logarithmique :

Nous avons réalisé cette fonction au moyen du montage proposé par Dobkin (1969). L'amplificateur est constitué par deux amplificateurs opérationnels (LF 357 H et LM 301 AH) et par une paire de transistors LM 394 qui améliore la dynamique de mesure par rapport à la paire originale LM 314. La variation de la tension de sortie, fonction du logarithme de la tension d'entrée est comprise entre + 2 et - 2 volts, le zéro correspond à un courant d'entrée de 10 microampères (Figure 22). Un amplificateur opérationnel inverseur permet d'obtenir, au moyen d'un réglage du gain et du décalage d'offset, une sortie calibrée. La dynamique de sortie est alors comprise entre 8 et 2 volts, valeurs auxquelles on fait correspondre les niveaux d'intensité de 80 à 20 dB. C'est avec cette sortie que nous utilisons l'intensimètre couplé avec un calculateur par l'intermédiaire d'un convertisseur analogique-numérique. Nous effectuons ainsi des mesures quantifiées et du traitement automatique sur les paramètres prosodiques (Teston et Rossi, 1977).

V - CONCLUSION

Nous venons de voir comment réaliser une mesure simple et précise de l'intensité objective des signaux de parole. Dès que l'on désire obtenir une évaluation de la sonie ou intensité effectivement perçue, la mesure se complique énormément. On peut corriger cette dernière en fonction des paramètres acoustiques du signal, au moyen de différentes méthodes.

Nous nous employons actuellement à réaliser dans cette perspective, un analyseur de sonie automatique. Plus complexe encore est l'influence de la durée des segments acoustiques et leur interaction sur le plan perceptif, au cours de l'évolution temporelle de la phrase. De nombreuses recherches restent à effectuer dans ce domaine à la lumière des résultats obtenus par les psychoacousticiens. Nous espérons, dans un proche avenir, pouvoir intégrer de tels résultats dans le calcul automatique de la sonie, tout en doutant de la possibilité de pouvoir un jour réaliser un sonomètre automatique qui tiendrait compte de tous les facteurs de correction que nous venons d'énumérer succinctement, sans oublier bien sûr ceux que nous n'avons pas mentionnés, ou dont on n'a pas encore pris connaissance.



$$\begin{aligned}
 A_{\text{RMS}} &= \frac{\pi}{2\sqrt{2}} \times A_{\text{moy}} = \frac{1}{\sqrt{2}} \times A_{\text{crête}} \\
 A_{\text{RMS}} &= 1,11 \times A_{\text{moy}} = 0,707 \times A_{\text{crête}} \\
 A_{\text{crête}} &= 1,57 \times A_{\text{moy}} = 1,414 \times A_{\text{RMS}} \\
 A_{\text{moy}} &= 0,9 \times A_{\text{RMS}} = 0,637 \times A_{\text{crête}}
 \end{aligned}$$

Figure 1
Définition des différentes amplitudes caractéristiques d'un signal sinusoïdal

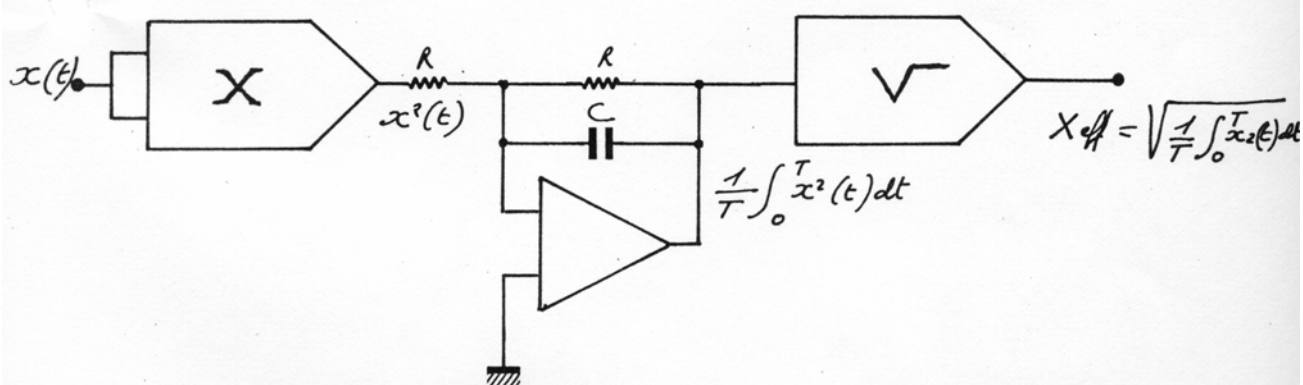


Figure 2
Schéma de principe d'un calculateur analogique de la valeur efficace d'un signal. Il est constitué par un multiplieur, un intégrateur et un extracteur de racine carrée

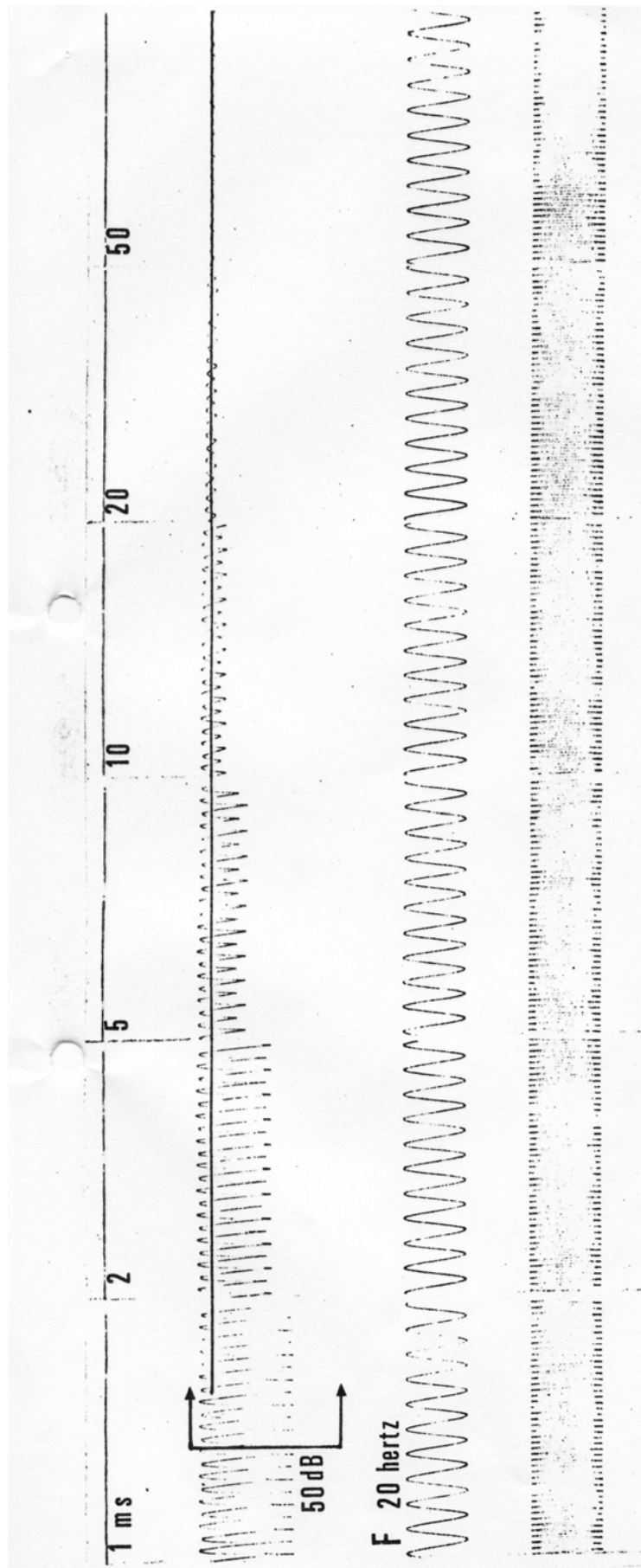


Figure 3

Mise en évidence des ondulations résiduelles (ripple) dans les cas de constantes de temps trop faibles par rapport à la période des signaux. Pour une fréquence de 20 Hertz, le critère d'équivalence nous indique une valeur de 25 millisecondes. On constate qu'avec 50 millisecondes, il n'existe quasiment pas d'ondulation résiduelle ; avec 20 millisecondes, l'ondulation par rapport à la valeur RMS est de $\pm 0,8$ dB. Plus on diminue la constante de temps, plus l'ondulation devient importante. On remarque sa dissymétrie par rapport à la valeur efficace. Le facteur de crête du signal n'étant pas très élevé (1,4) on a intérêt à prendre la valeur crête de la valeur efficace, on minimise ainsi l'erreur de mesure, qui est ici de +3 dB avec une constante de temps de 5 millisecondes. Les ondulations ne sont pas constantes en amplitude dans un faible rapport, ceci est dû au fait que le générateur sinusoïdal que nous utilisons possède une légère tension continue de décalage (offset).

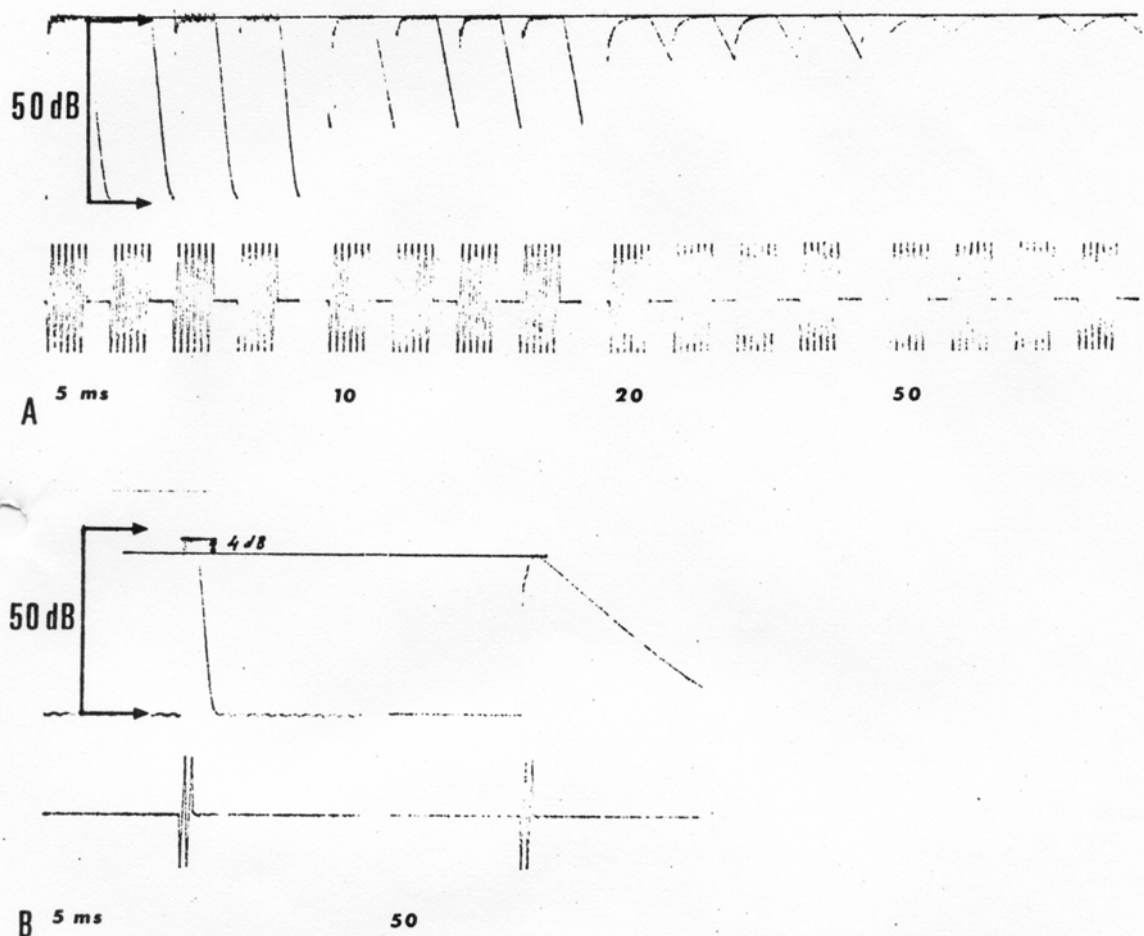


Figure 4

- A) Mise en évidence de l'influence de la constante de temps d'intégration sur la mesure de l'amplitude de bouffées de signaux (tone burst). Pour de fortes valeurs d'intégration, la valeur de la mesure n'est pas perturbée mais l'intensimètre "voit" du signal partout.
- B) Mise en évidence des erreurs de mesure sur les transitoires dues à une trop forte constante de temps d'intégration. Dans cet exemple, le transitaire n'en est d'ailleurs pas un, il dure 20 millisecondes et est constitué par 2 périodes de signaux triangulaires à 100 Hertz. Le résultat est malgré tout significatif.

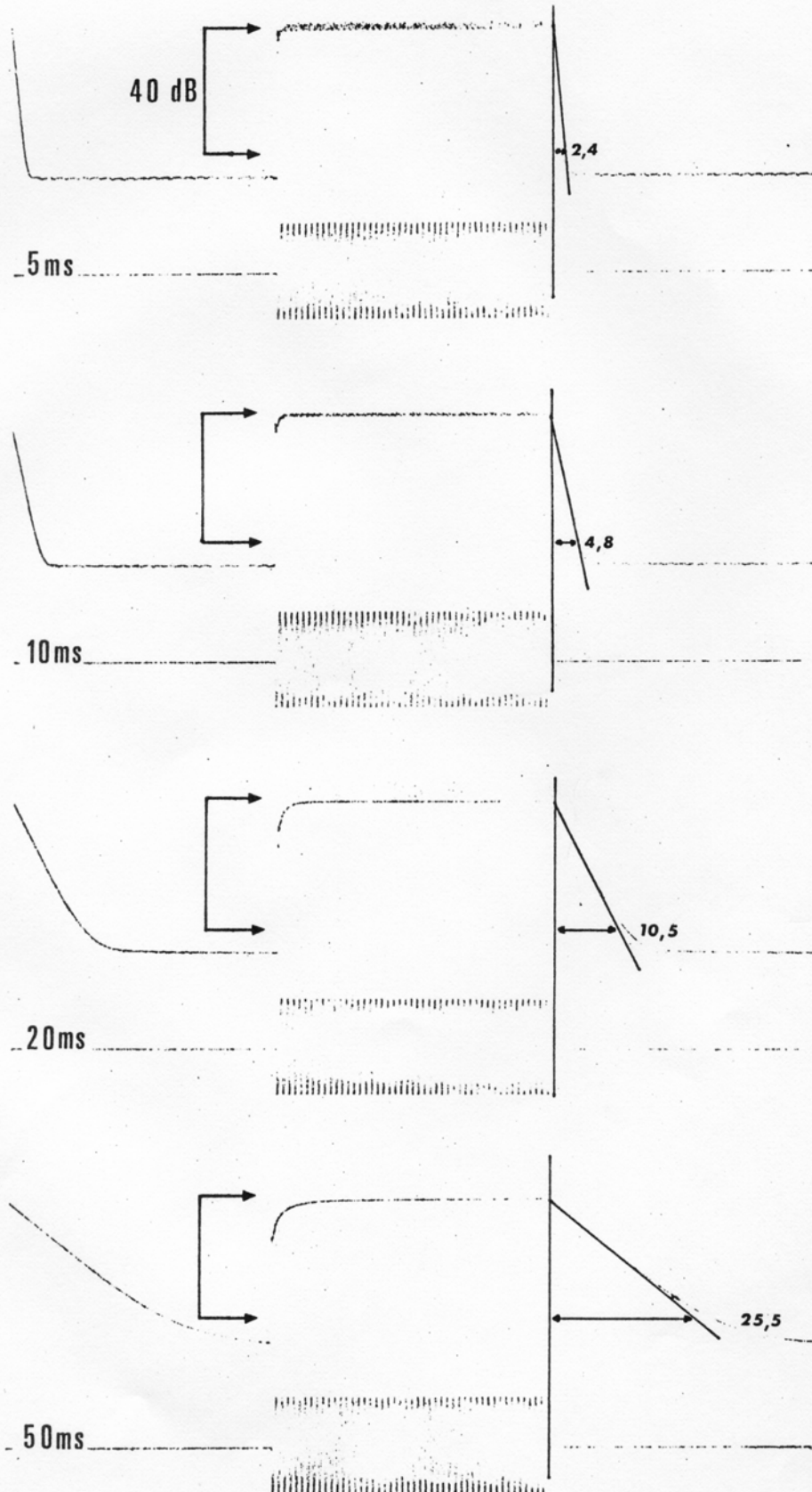


Figure 5

Mise en évidence des temps d'établissement et de descente d'une intégration moyenne pondérée exponentielle. Nous obtenons par le calcul respectivement les valeurs de : 5,21 - 10,48 - 22,59 et 54,38 pour les constantes de temps d'intégration.

Figure 6

Mise en évidence des erreurs de mesure effectuées en fonction de l'ondulation résiduelle. Celle-ci est liée à la nature du signal dont on mesure l'intensité. Nous avons vu à la figure 3 ce qu'il en était pour un signal sinusoïdal.

Nous avons mesuré les niveaux à partir de la constante de temps de 50 millisecondes pour avoir la mesure la plus précise par rapport à la durée des voyelles. Si l'on prend les valeurs crête de la valeur efficace, nous avons :

- Pour la voyelle A : Même valeur pour 5, 10, 20 et 50 ms.
- Pour la voyelle I : + 1 dB avec 5 ms, + 0,5 dB avec 10 ms, 0 dB avec 20 et 50 ms.
- Pour la voyelle OU : + 2 dB avec 5 ms, + 1,5 dB avec 10 ms, + 0,5 dB avec 20 ms et 0 dB avec 50 ms.

Ces différences sont dues au facteur de crête des signaux des différentes voyelles plus particulièrement à la dissymétrie par rapport au zéro, et à leur énergie dans le bas du spectre. D'une manière générale nous préconisons de mesurer la valeur efficace aux crêtes lorsque l'on est en présence d'une ondulation résiduelle. Cette mesure a pour avantages :

- de pouvoir échantillonner la courbe d'intensité au moyen de la fréquence fondamentale à chaque coup de glotte.
- de réaliser la mesure la plus facile à mettre en oeuvre sans avoir un taux d'erreur important.

Certains auteurs prennent le milieu des ondulations, l'erreur est du même ordre et le milieu de l'ondulation plus difficile à évaluer. De toute façon la valeur efficace ne se situe pas au milieu de l'ondulation résiduelle.

A

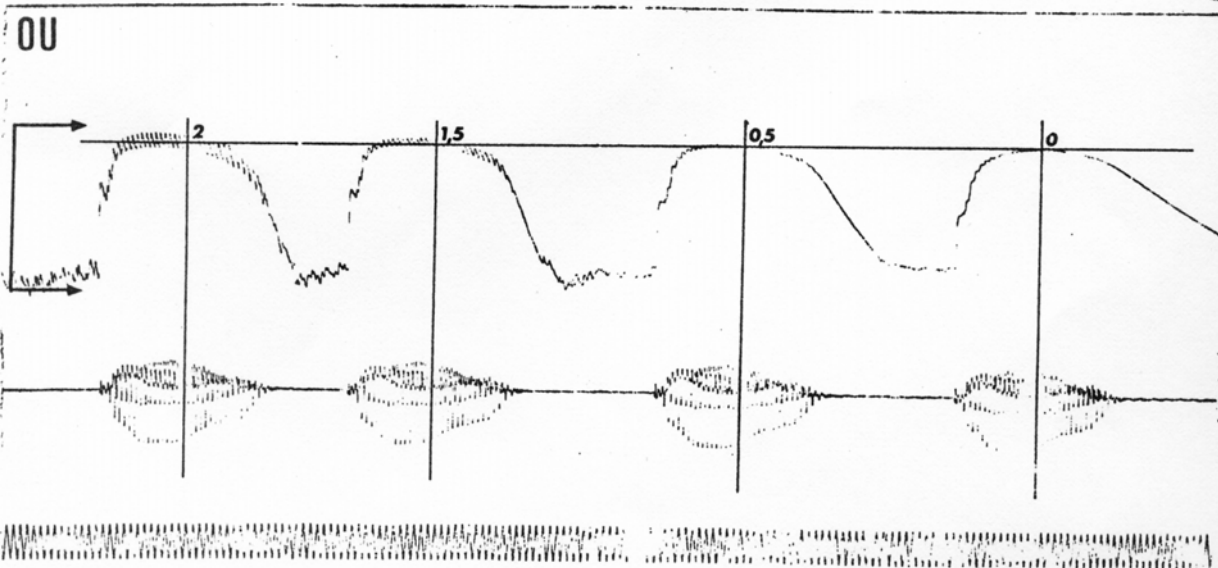
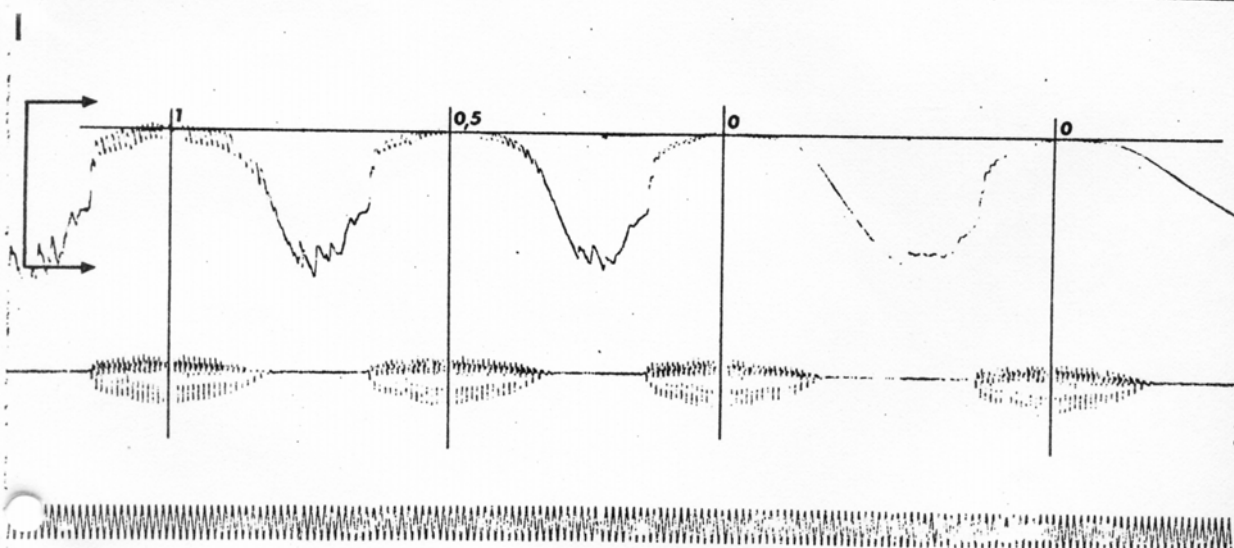
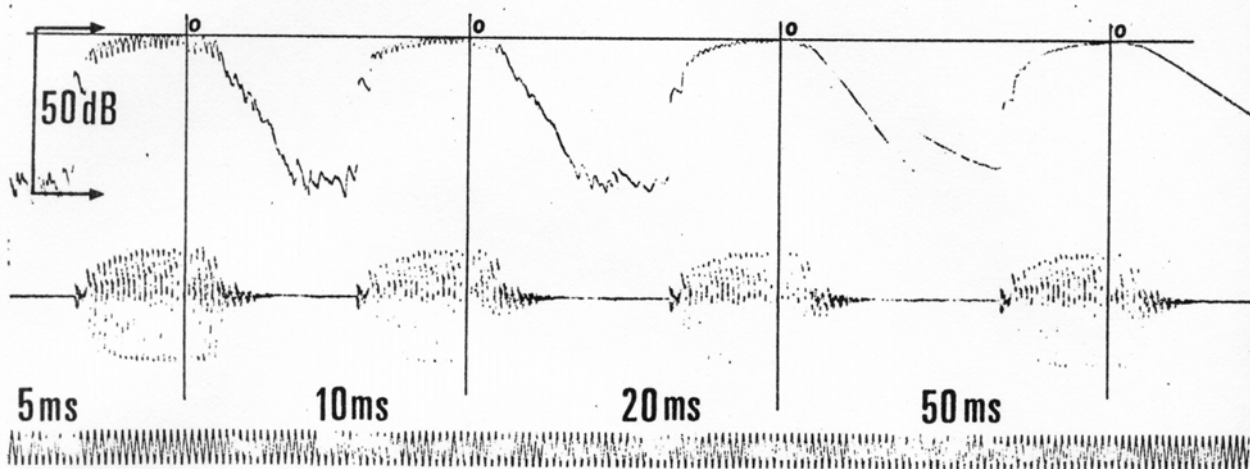


Figure 7

Même procédure que précédemment mais avec deux séquences VCV : ASA et ACHA. On remarque le fait rapporté dans la figure 6 au sujet de la voyelle A que l'on retrouve au même niveau d'intensité dans la gamme des constantes de temps que nous utilisons, ceci pour les voyelles terminales qui ont une durée supérieure à 150 ms. Ce n'est pas vrai par contre pour les voyelles initiales beaucoup plus courtes (35 ms pour ASA et 50 ms pour ACHA) ou le niveau d'intensité décroît avec l'augmentation de la constante de temps. On arrive à - 3 dB avec 50 ms pour la voyelle la plus courte ; l'initiale de ASA.

On remarque que l'intensité des consonnes constrictives est très importante et arrive presque dans ACHA au niveau de l'intensité des voyelles. Ceci peut paraître curieux a priori, mais dans certains cas les constrictives ont parfois plus d'énergie que les voyelles adjacentes, contrairement à ce que l'on imagine généralement et ceci pour trois raisons essentielles :

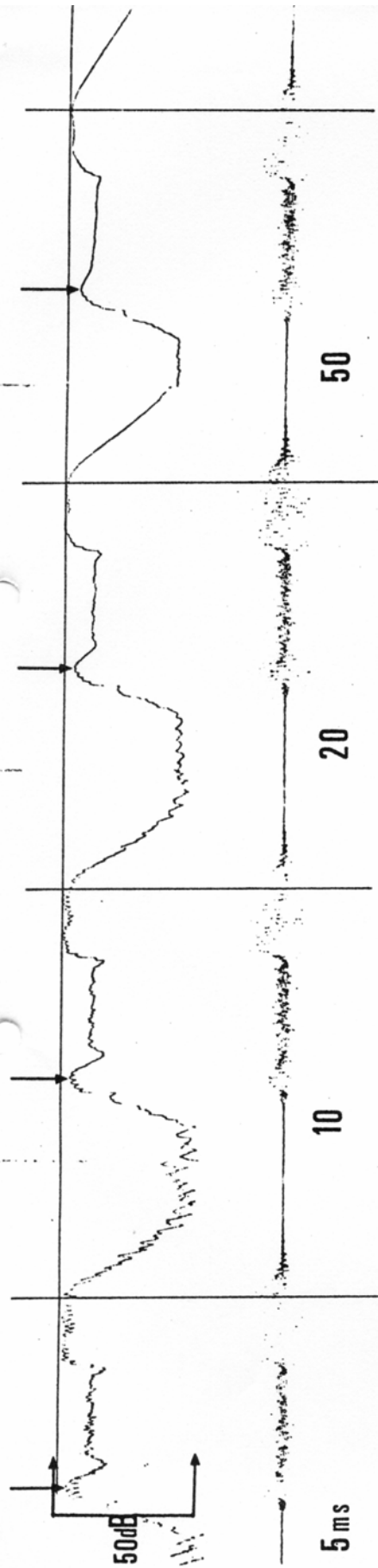
- La première est que ce type de signaux est voisés pour beaucoup de raisons que nous ne développons pas.

- La seconde est que les oscillogrammes de parole, sont visualisés au moyen d'enregistreurs oscillographiques dont la bande passante en fréquence est limitée en dessous de la valeur de 1 KHz (Tous les documents de nos figures). Dans de telles conditions de restriction spectrale, l'amplitude des consonnes constrictive est très réduite.

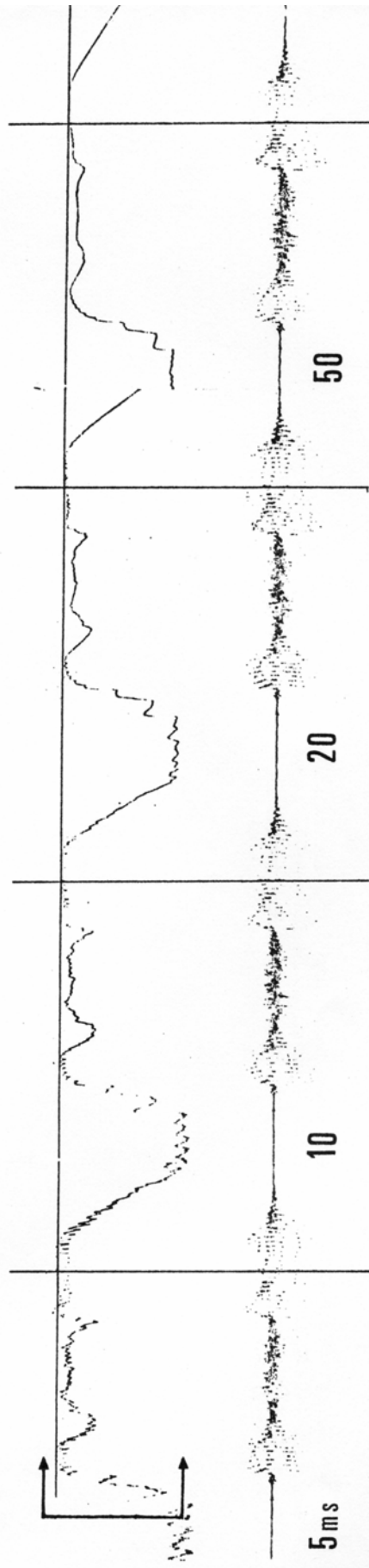
- La troisième enfin, trouve sa source dans le fait que certains intensimètres ne disposent pas de détecteurs de valeur efficace capable de détecter de manière satisfaisante des bruits à temps de montée rapide malgré leur bande passante élevée (15 KHz).

Il est important de noter que l'on doit prendre garde aux conditions d'enregistrement qui peuvent jouer un rôle non négligeable dans l'intensité relative des consonnes constrictives.

ASA



ACHA



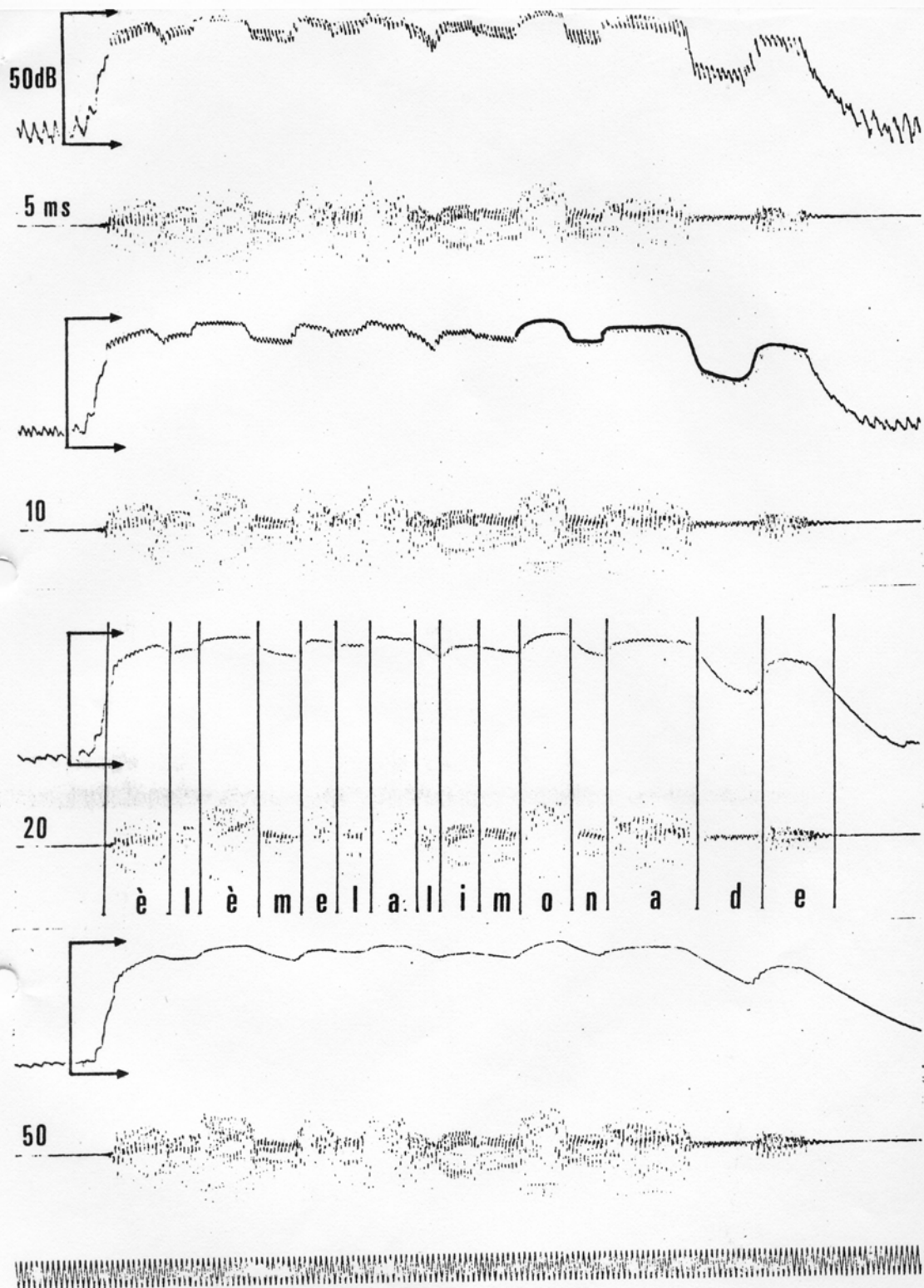


Figure 8
"ELLE AIME LA LIMONADE"

Cette phrase représente un continuum voisé. Les variations d'amplitude de la courbe d'intensité sont relativement faibles. La constante de temps de 10 millisecondes donne une bonne définition de l'enveloppe. La constante de temps de 20 ms commence à lisser légèrement les variations d'amplitude. La courbe enveloppe des ondulations résiduelles que nous avons tracées sur la fin de la phrase avec une constante de temps de 10 ms donne une lecture de l'intensité avec une erreur maximale de + 1 dB.

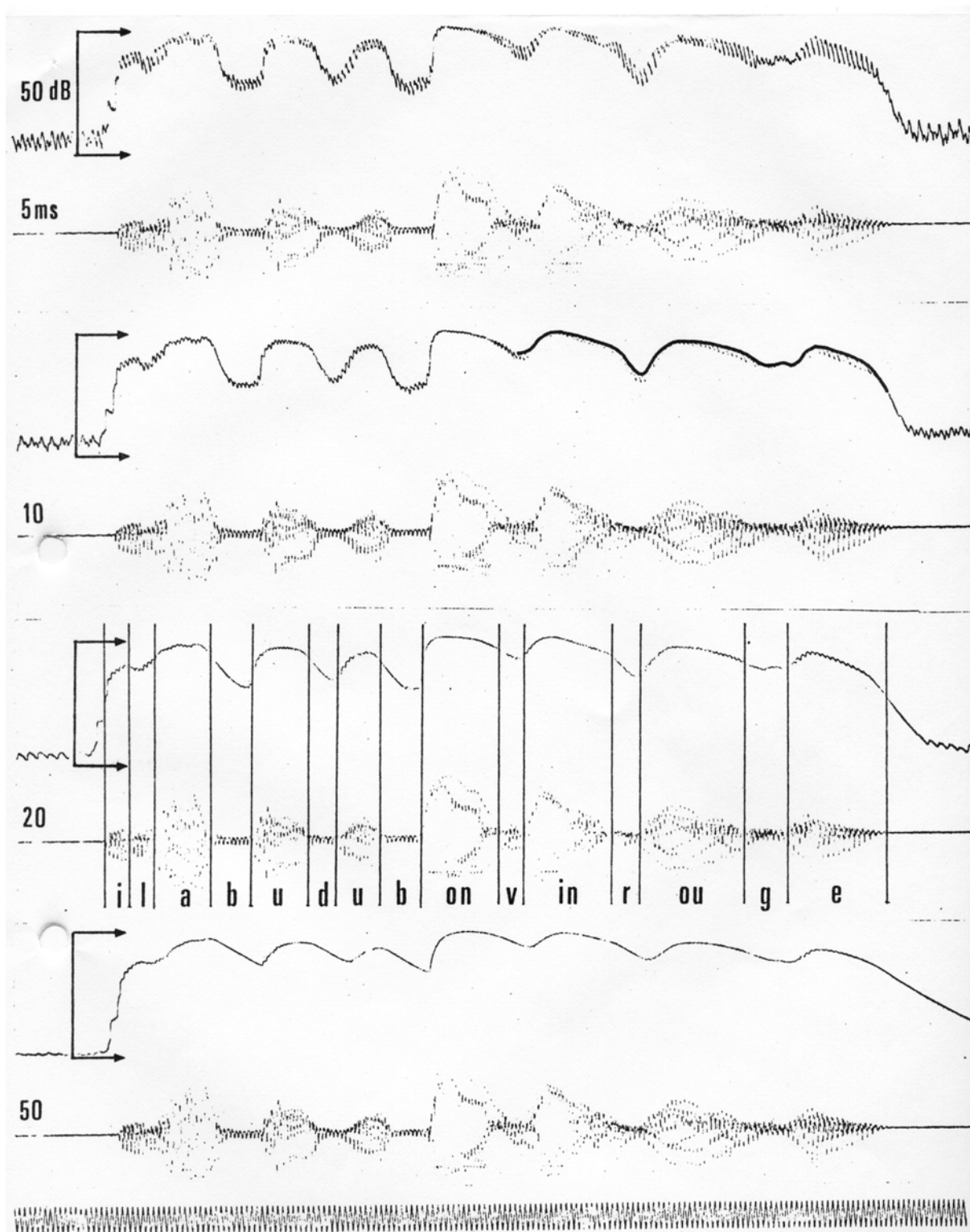


Figure 9

"IL A BU DU BON VIN ROUGE"

Phrase représentant comme précédemment un continuum voisé mais avec dans ce cas des variations d'amplitude beaucoup plus importantes, dues à la présence de plusieurs consonnes (B et D en particulier). La constante de temps de 10 millisecondes offre le meilleur compromis. Même remarque que pour la Figure 8 au sujet de la précision de la mesure.

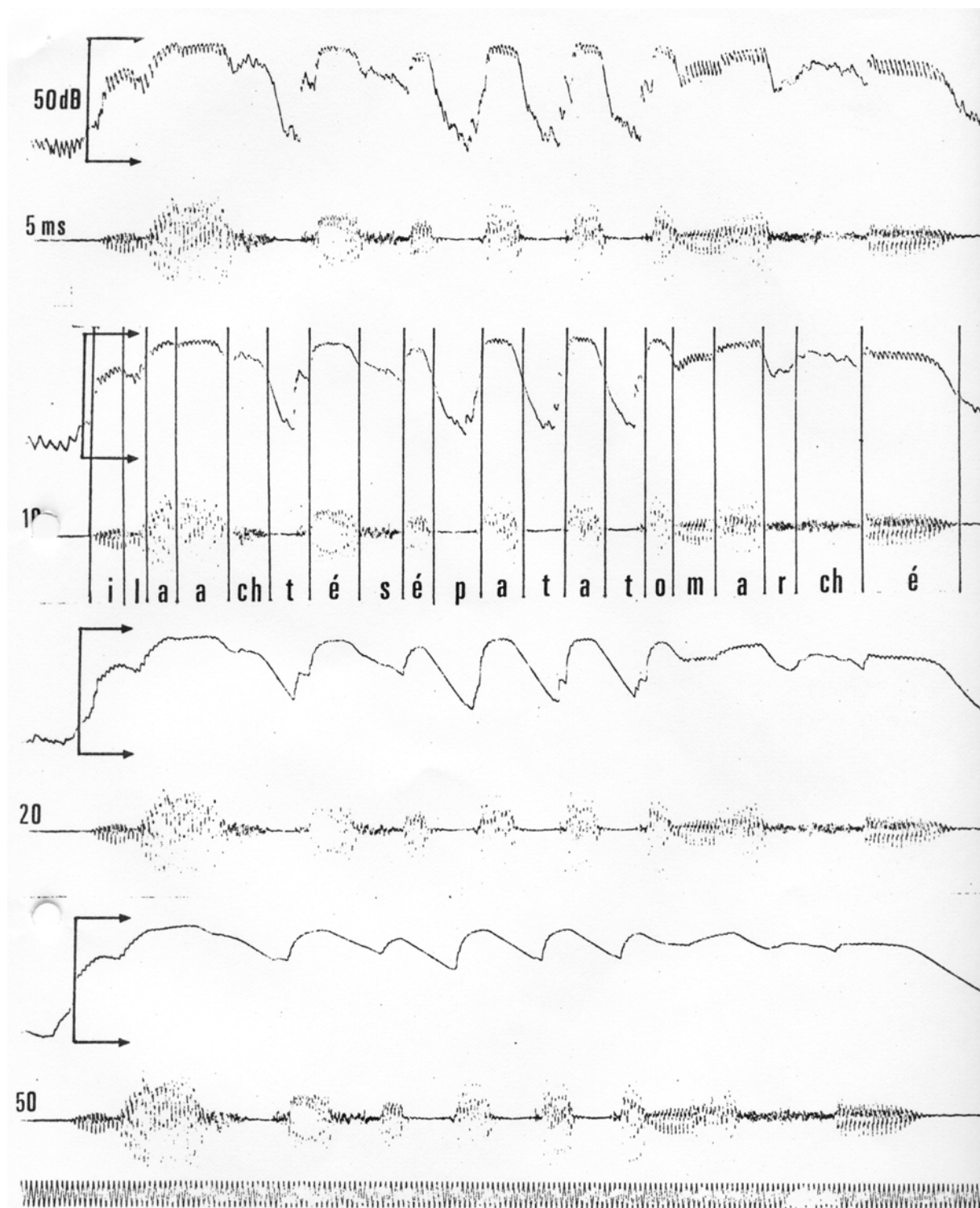


Figure 10

"IL A ACHETE SES PATATES AU MARCHE"

Exemple de phrase particulièrement découpée par de nombreuses consonnes occlusives et constrictives. La constante de temps de 10 millisecondes s'impose instantanément. On note la présence de nombreux transitoires d'explosions de consonnes occlusives. Ils sont mesurés avec une erreur inférieure à 1 dB par rapport à la constante de temps de 5 ms. La précision des mesures sur les autres segments est du même ordre que précédemment.

L'effet de lissage des variations d'intensité est particulièrement visible lorsque l'on utilise des constantes de temps longues (20 et 50 ms).

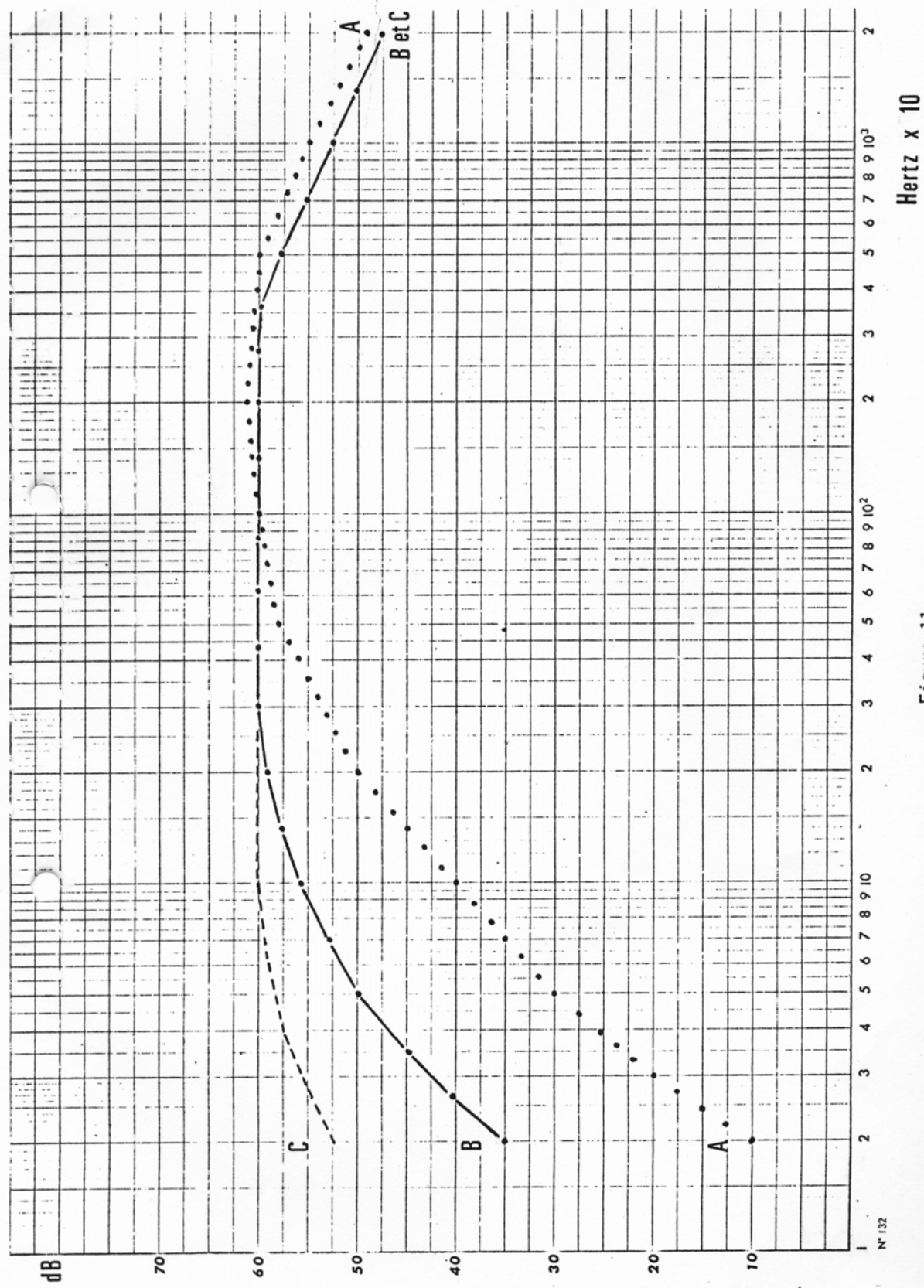


Figure 11

Courbes de pondérations subjectives normalisées A, B et C, qui correspondent respectivement aux niveaux de 40, 70 et 100 PHONES.

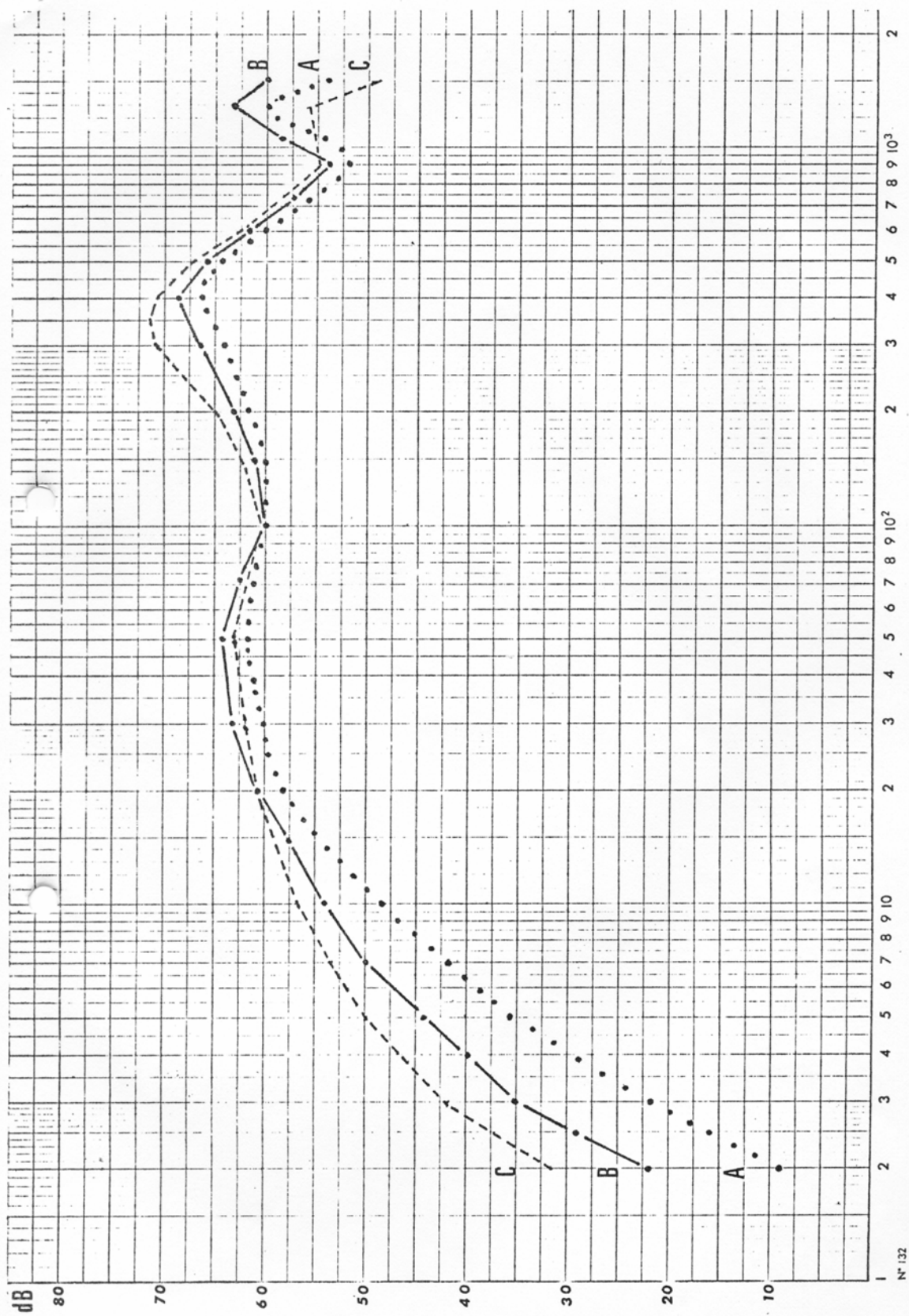


Figure 12

Courbes de pondérations subjectives effectivement mesurées à partir des courbes isosoniques NF S 30-003 pour les niveaux de 40, 70 et 100 PHONES. On remarque les différences importantes entre ces mesures et les courbes normalisées de la figure précédente.

Hertz x 10³

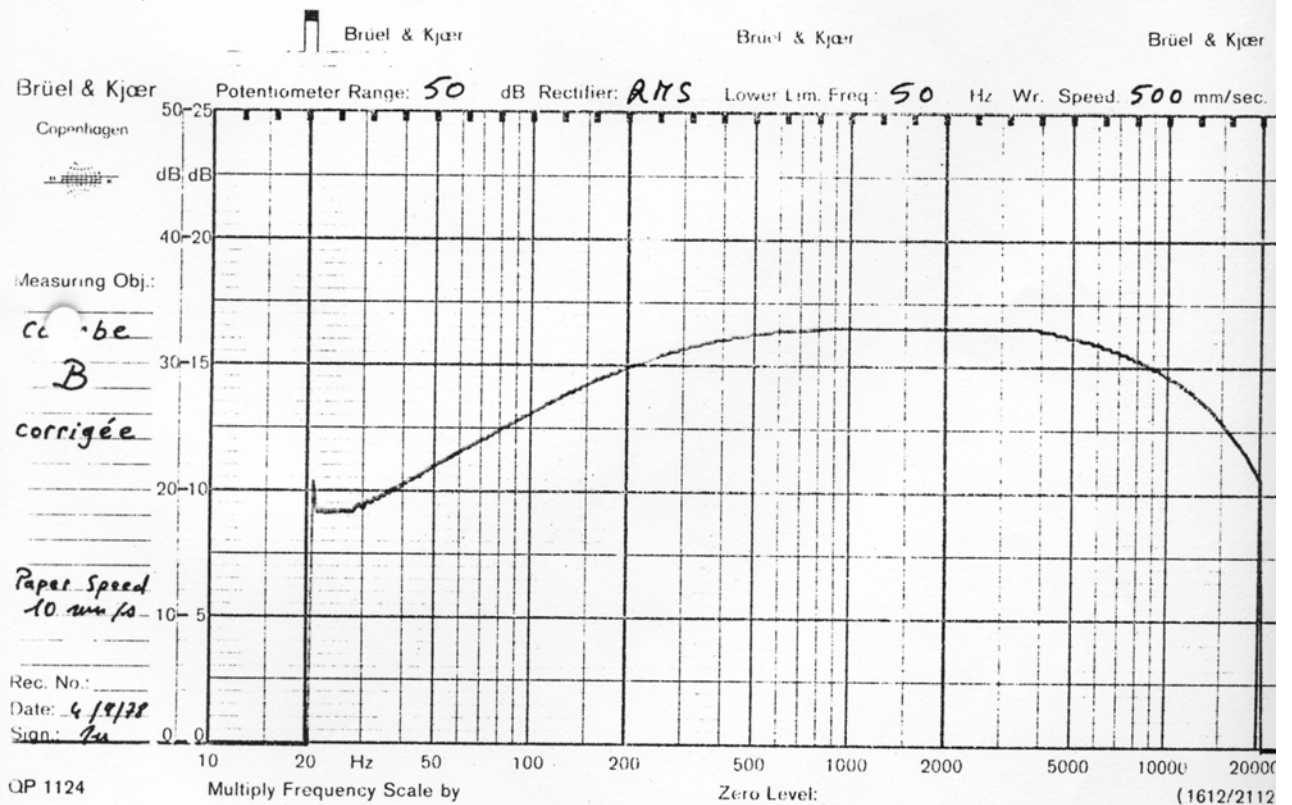
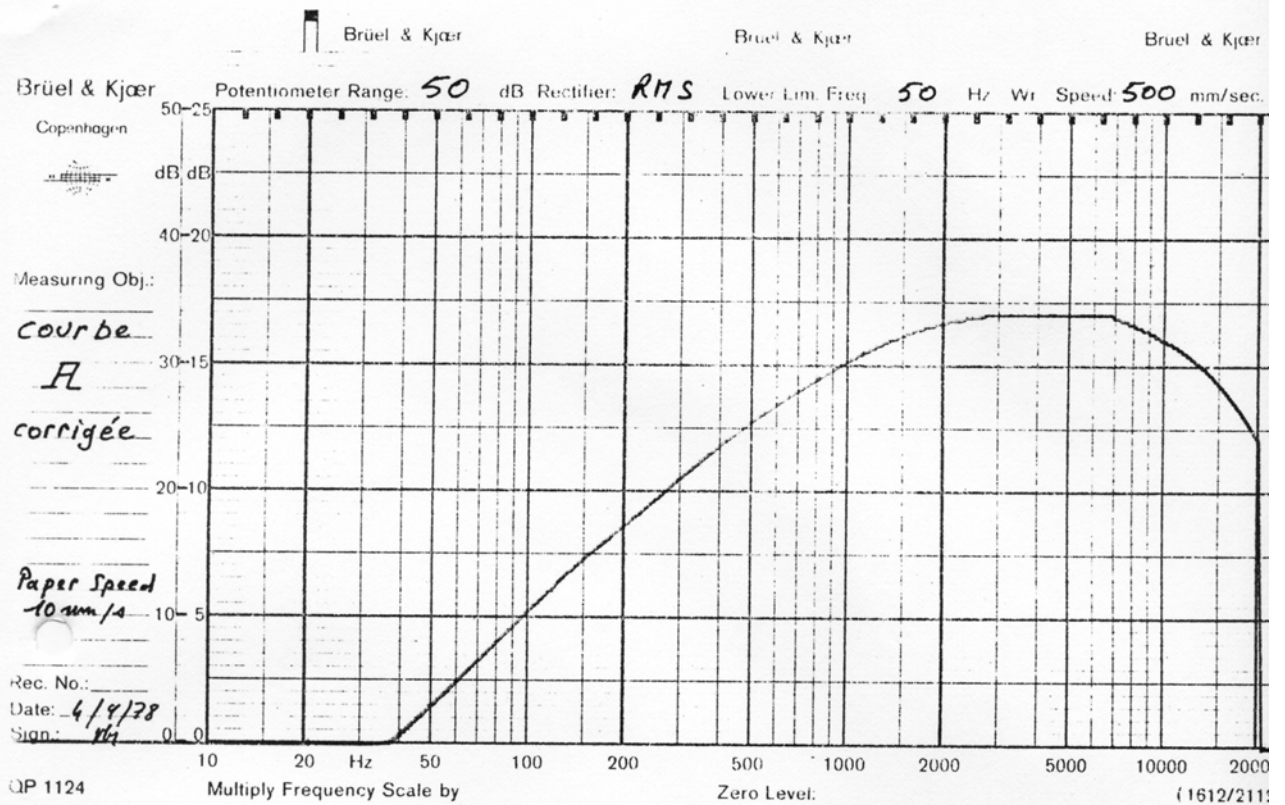


Figure 13 a

Courbes de réponse en fréquence des pondérations A et B de l'intensimètre.

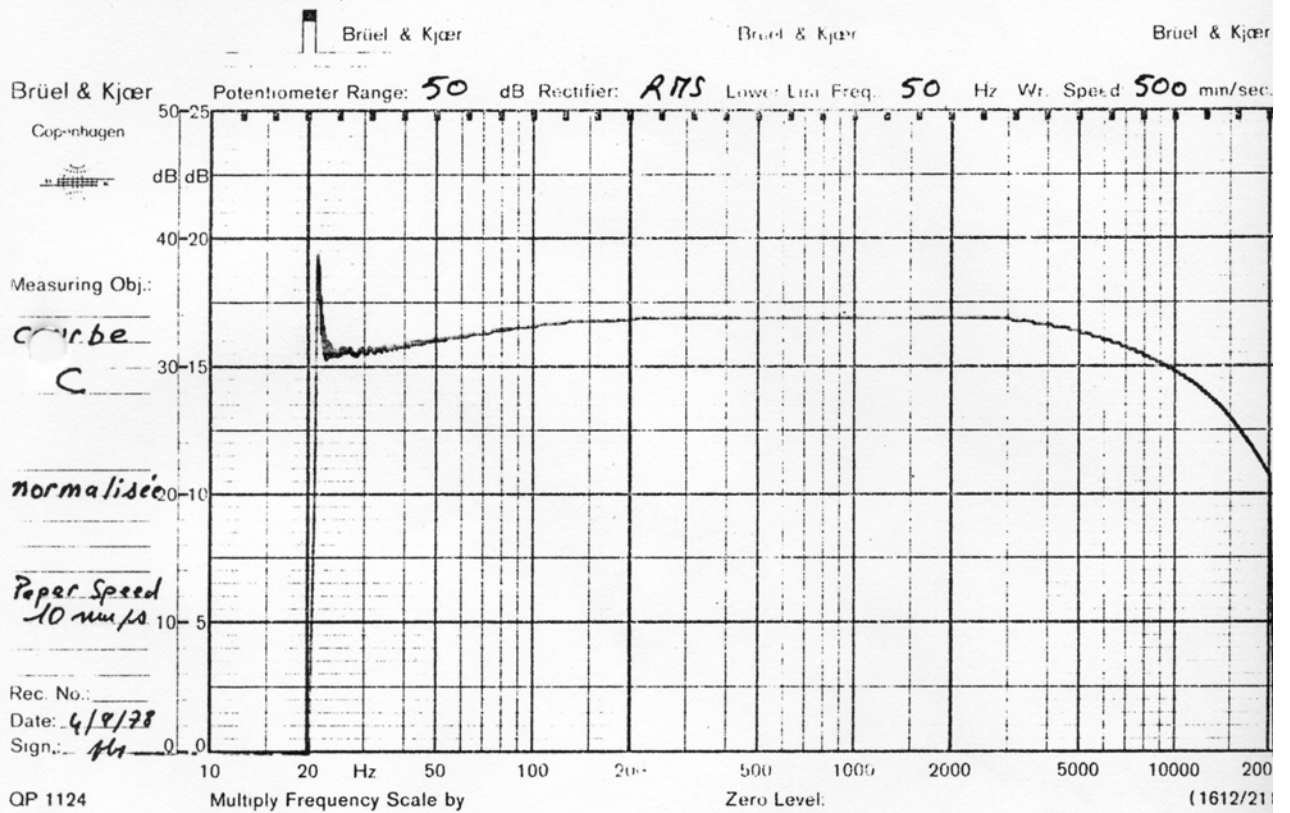


Figure 13 b

Courbe de réponse en fréquence de la pondération C de l'intensimètre.

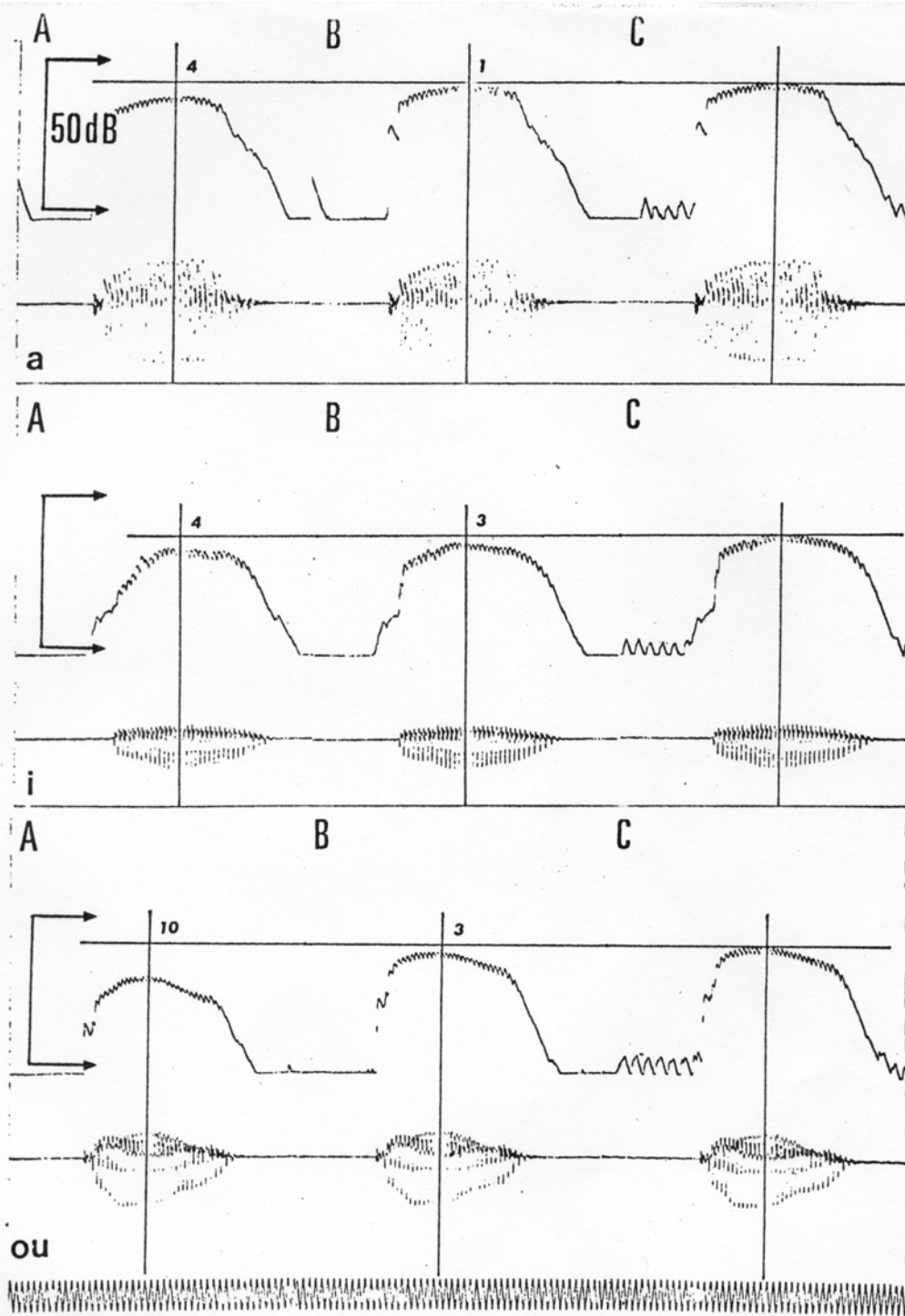


Figure 14

Comparaison des mesures de l'intensité effectuée au moyen des trois courbes de pondération A, B et C pour les voyelles a, i et ou.

Par rapport à la courbe la plus linéaire (la courbe C) on remarque :

Pour la voyelle a : - 1 dB avec la courbe B

" " " - 4 dB " " " A

" " " i : - 3 dB " " " B

" " " - 4 dB " " " A

" " " ou : - 3 dB " " " B

" " " - 10 dB " " " A

Ces résultats sont cohérents avec la structure acoustique des voyelles. L'atténuation en courbe A est plus forte pour les voyelles dont le maximum d'énergie est concentré dans le bas du spectre. (1er et 2ème formants).

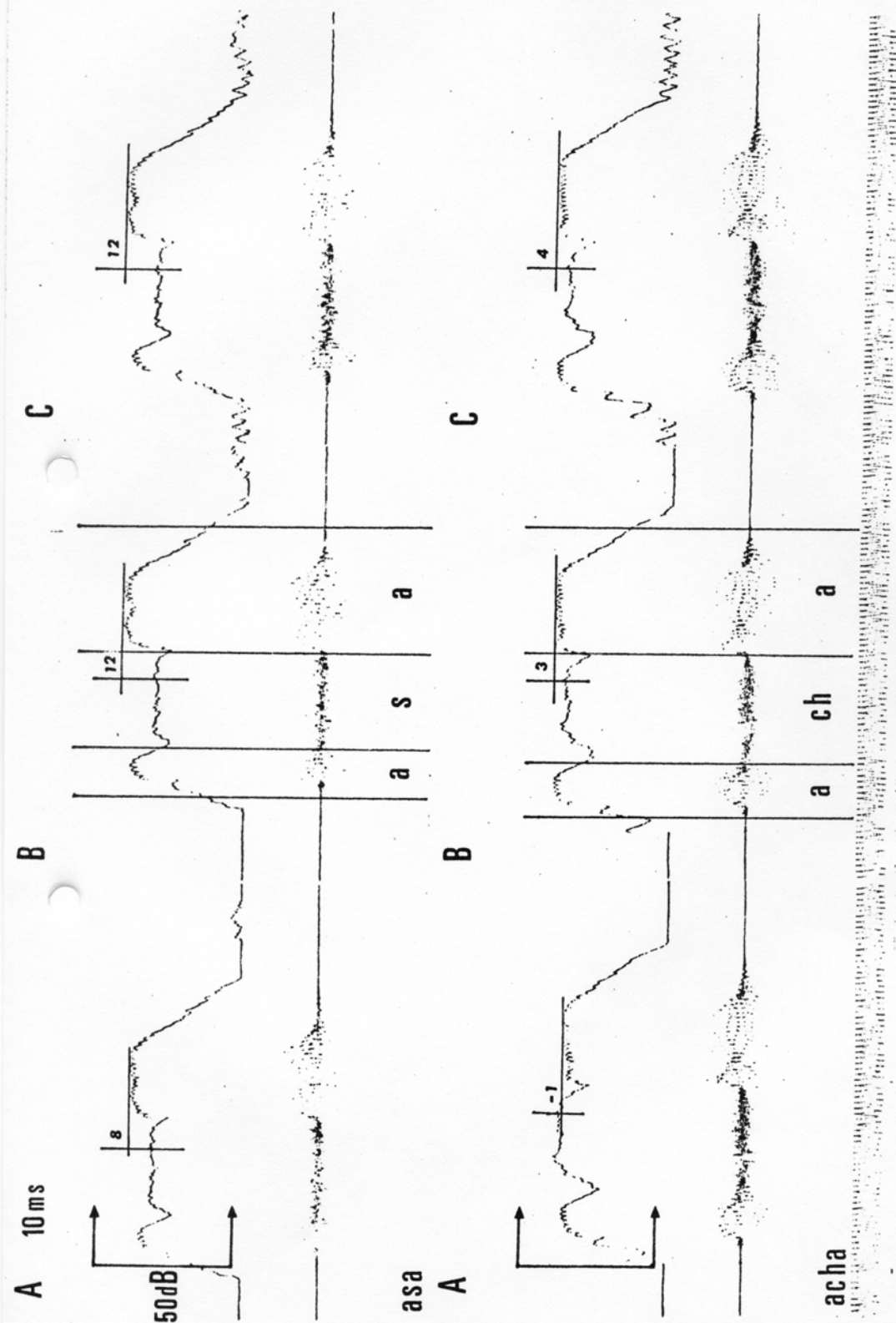


Figure 15

Mise en évidence de l'influence des pondérations A, B et C sur la mesure des consonnes constrictives. Leur structure acoustique étant différente de celle des voyelles (maximum d'énergie situé plus haut dans le spectre) elles sont moins atténuées. Il en résulte une relative amplification du niveau d'intensité des consonnes constrictives par rapport aux voyelles. Dans cet exemple, la voyelle a encadre les deux consonnes s et ch. Avec la courbe A, la différence d'intensité entre la voyelle finale et la consonne est de - 4 dB pour s et - 5 dB pour ch, par rapport à la courbe C que nous prenons comme référence.

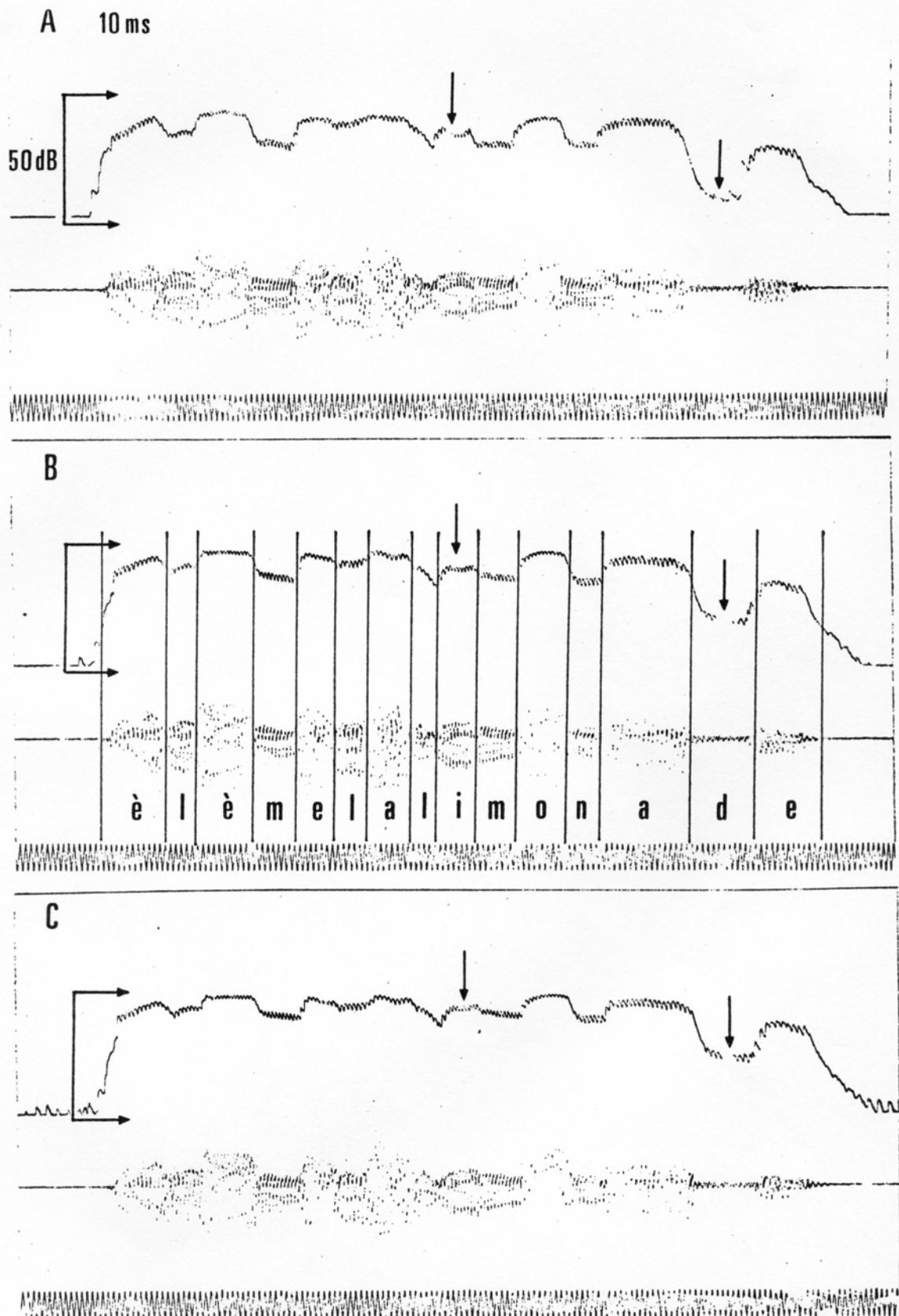


Figure 16

"ELLE AIME LA LIMONADE"

Phrase mesurée au moyen des pondérations A, B et C. On remarque l'influence des pondérations surtout avec la courbe A sur les différences de niveaux relatifs entre les voyelles et les consonnes.

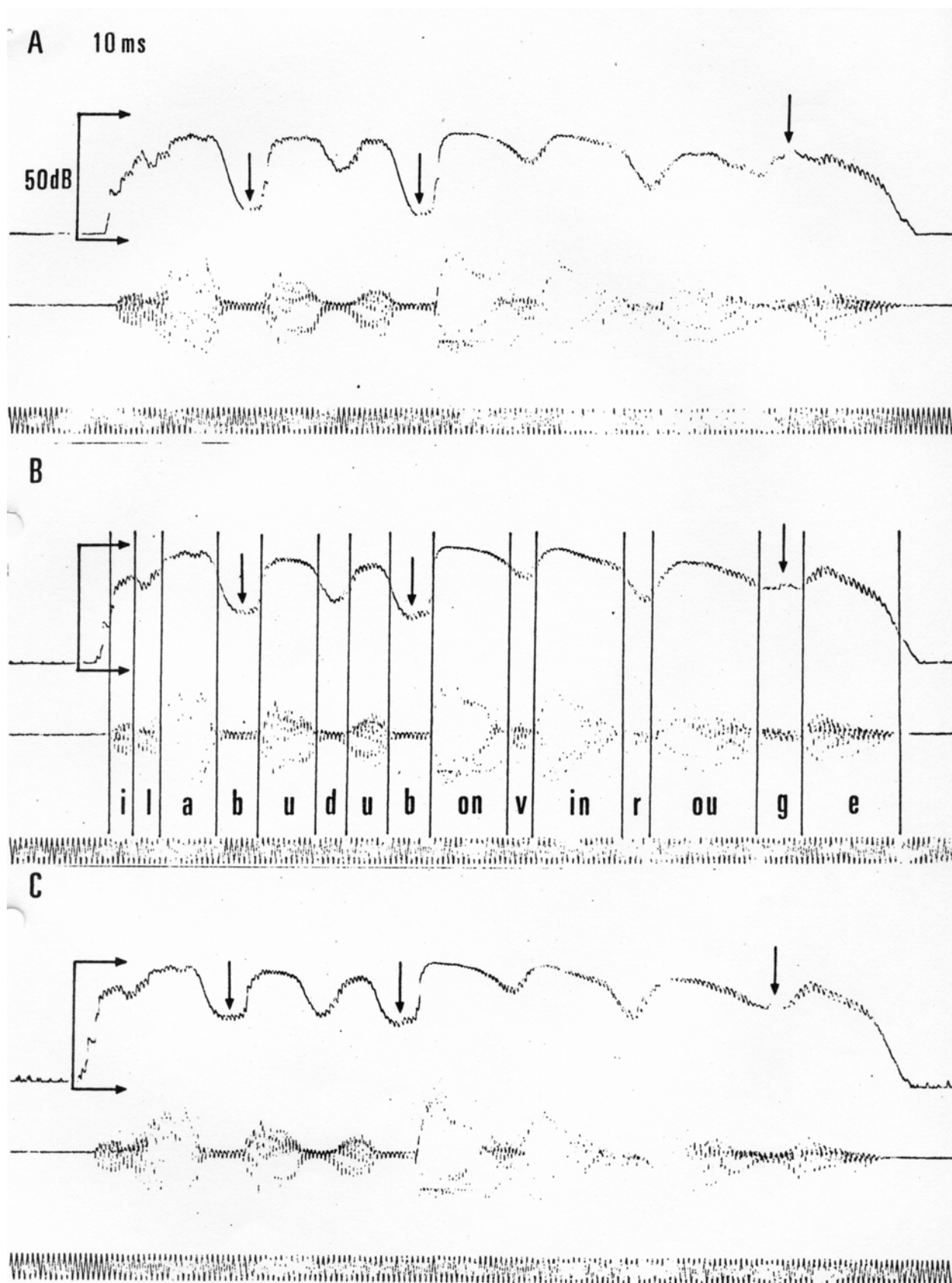


Figure 17

"IL BU DU BON VIN ROUGE"

Même procédure et même remarque que pour la figure précédente. Les différences de niveaux sont encore plus marquées notamment sur la consonne voisée "j".

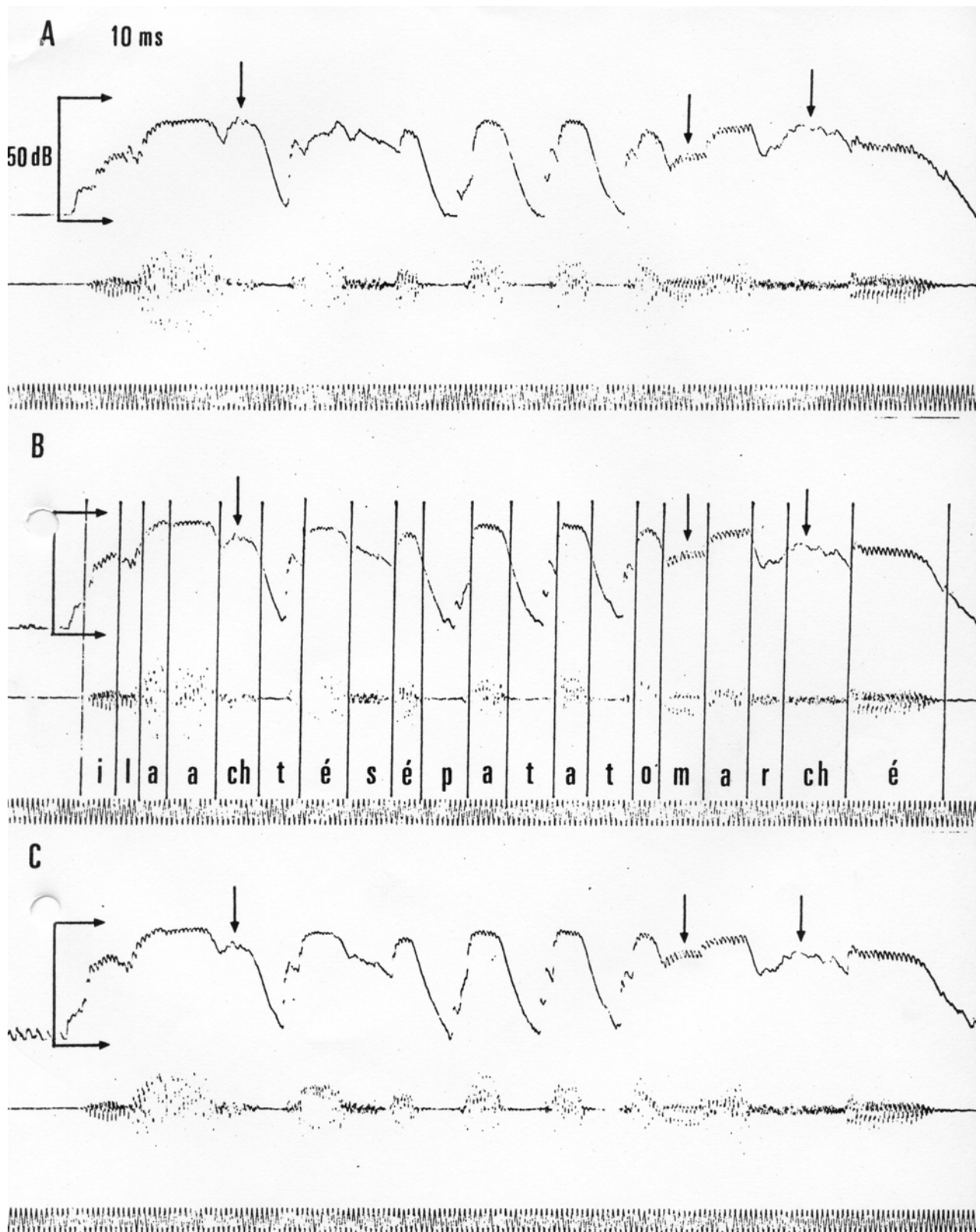


Figure 18

"IL A ACHETE SES PATATES AU MARCHE"

Mêmes remarques que pour les exemples précédents avec cette fois des consonnes constrictives.

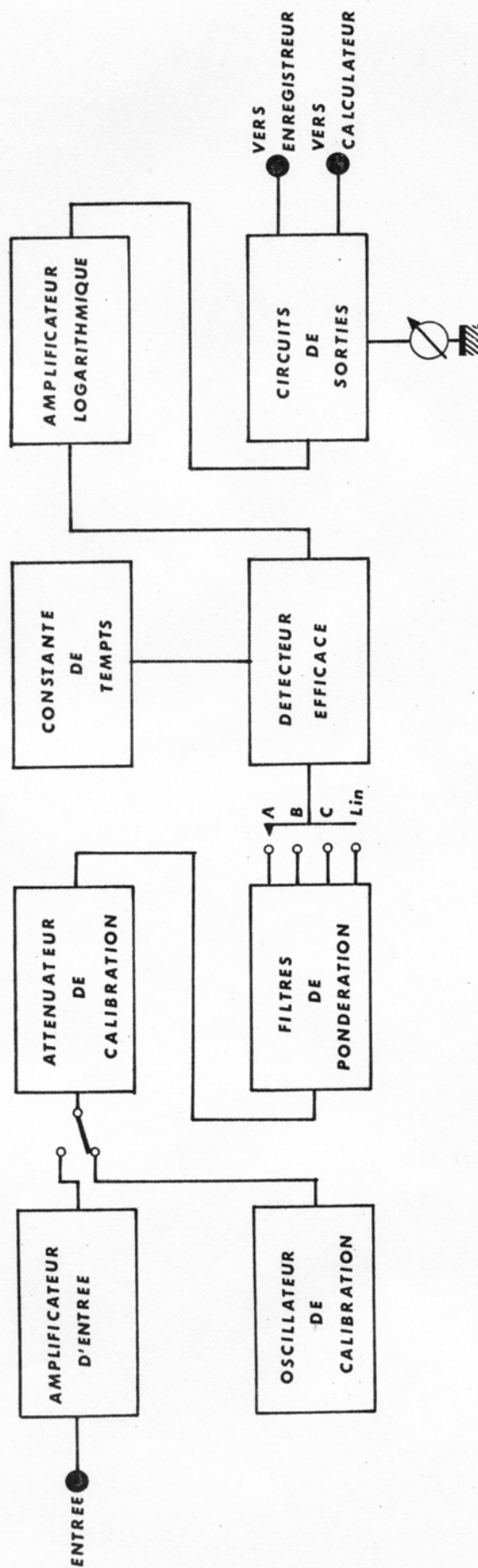


Figure 19
Schéma de principe de l'intensimètre

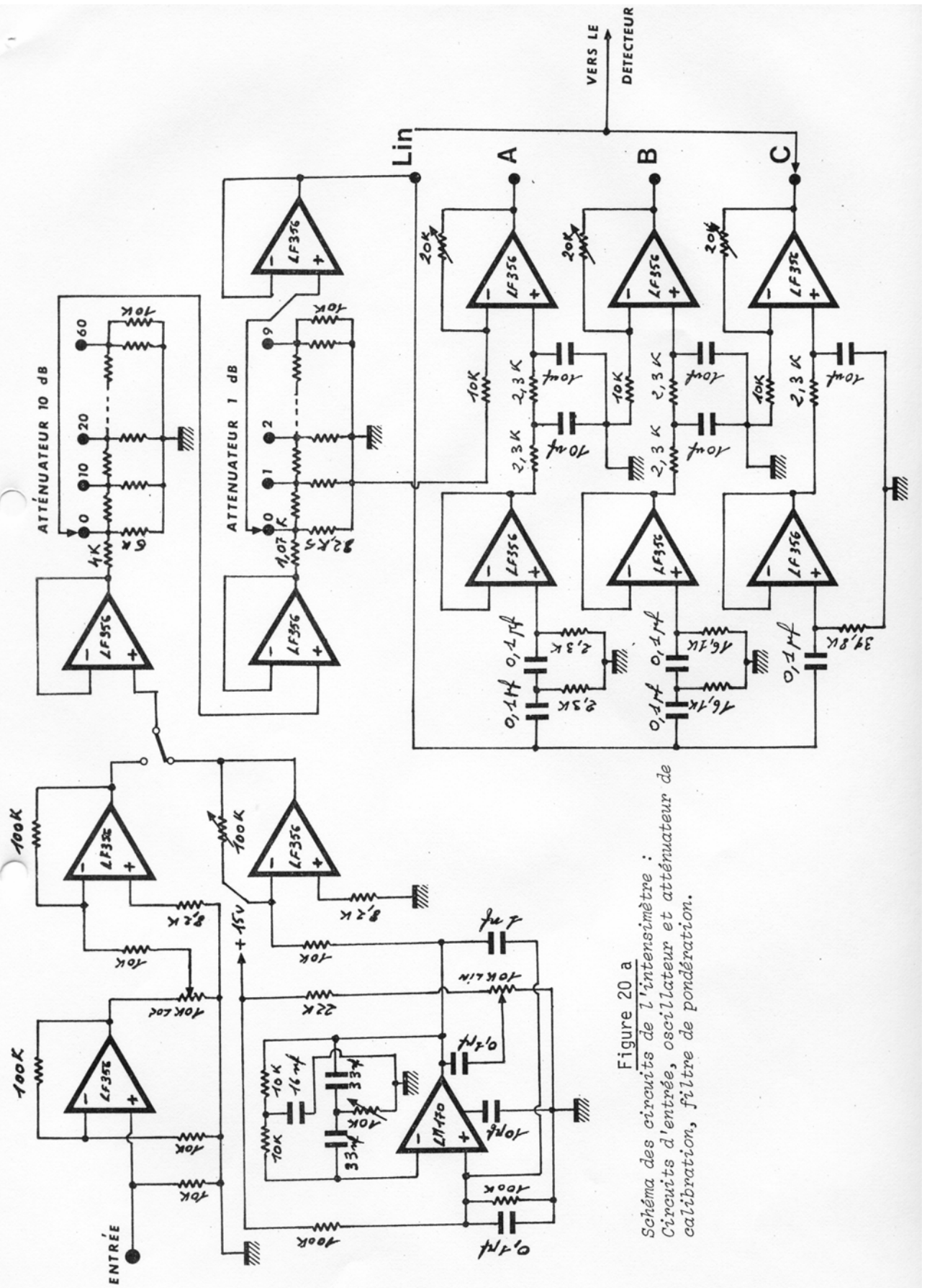


Figure 20 a
Schéma des circuits de l'intensimètre :
Circuits d'entrée, oscillateur et atténuateur de
calibration, filtre de pondération.

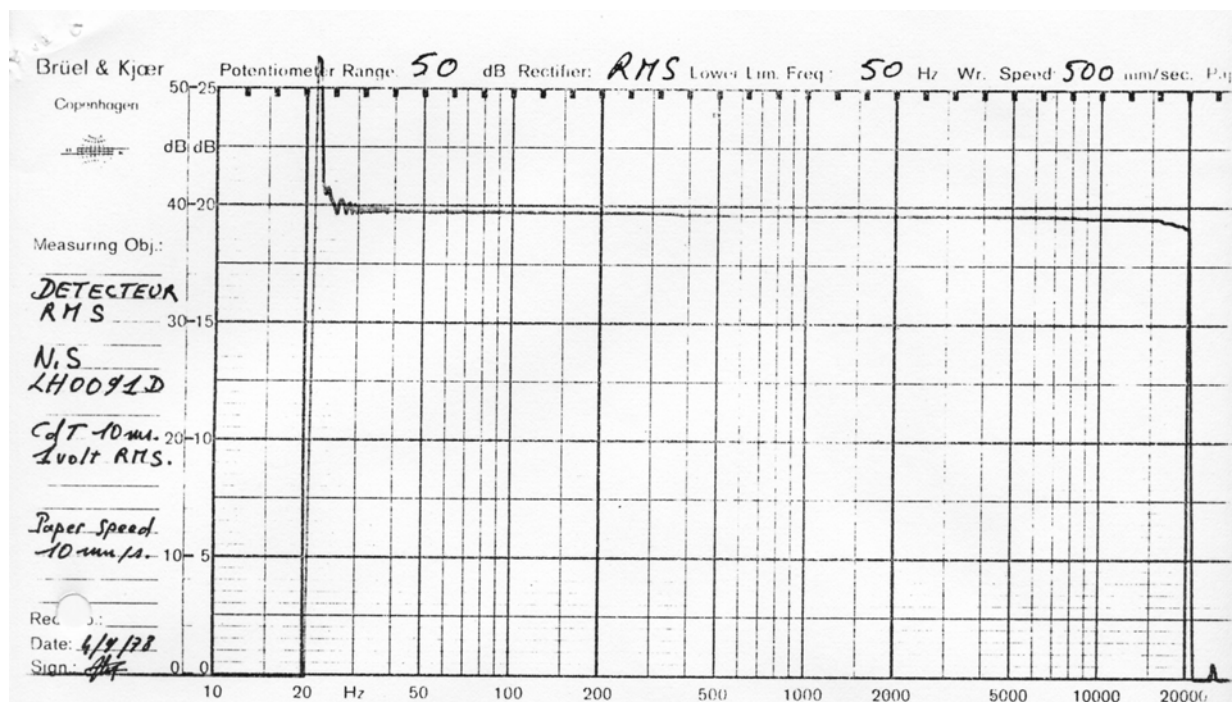


Figure 21

Courbe de réponse en fréquence du détecteur de valeur efficace LH0091CD.

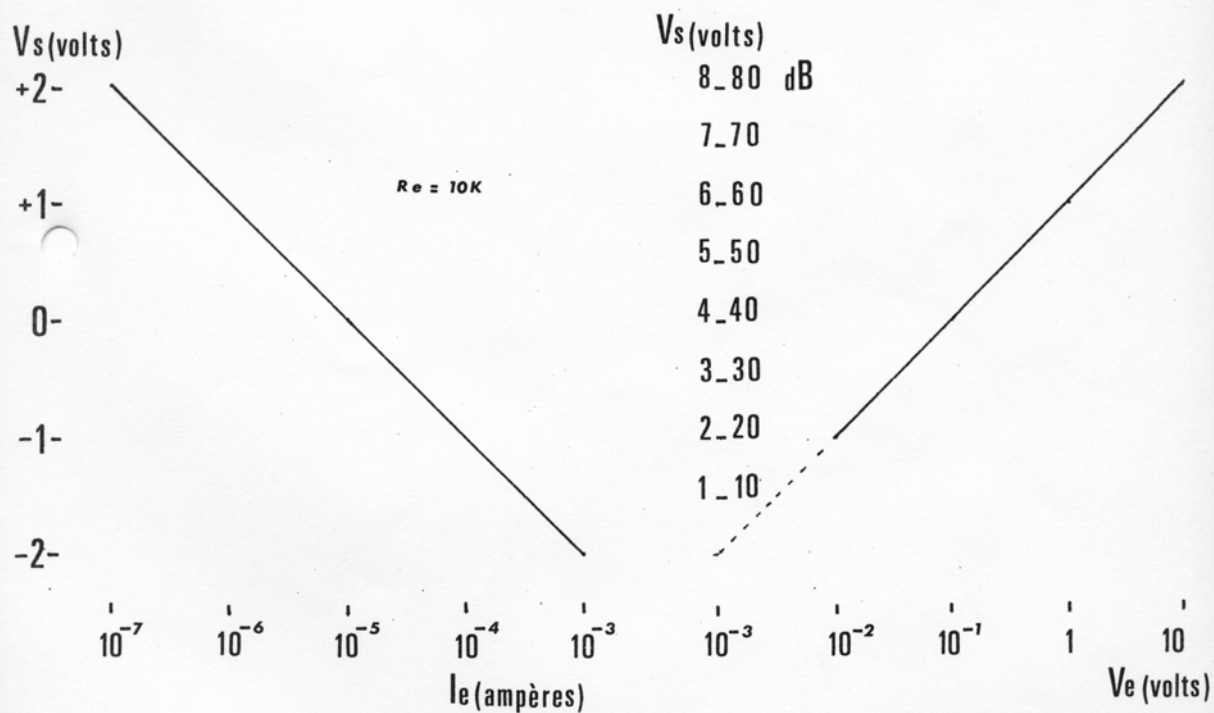


Figure 22

- A) Dynamique de l'amplificateur logarithmique de l'intensimètre.
B) Dynamique du signal de sortie de l'intensimètre. On a fait correspondre la tension de sortie avec un niveau en décibels, pour réaliser des mesures d'intensité au moyen d'un voltmètre.

BIBLIOGRAPHIE

- BAUD, J.C. (1978)
"Mesure des valeurs efficaces avec les multimètres numériques et analogiques", Electronique application N° 5, pp. 137-142.
- BEKESY, VON, G. (1960)
Experiments in hearing, Mac Graw Hill, New-York, pp. 321-330.
- BOE, L.J. et RAKOTOFIRINGA, H. (1971)
"Exigences, réalisation et limite d'un appareillage destiné à l'étude de l'intensité et de la hauteur d'un signal acoustique." Revue d'acoustique, N° 14, pp; 104-113.
- BRIE, C. (1971)
"Les quadrateurs analogiques à réseau de diodes", Mesure, vol. Mai, pp. 115-124.
- DOBKIN, R.C. (1969)
"Logarithmic converters", Linear Application I an 30, National semiconductor Corporation, Santa Clara.
- DOBKIN, R.C. (1973)
"True RMS, detector", Linear Application 2, LB 25, National Semiconductor Corporation, Santa Clara.
- FLANAGAN, J.L. (1960)
"Analog measurement of sound radiation from the mouth". J.A.S.A., vol.32, N° 12, pp. 1613-1620.
- FLETCHER, H. et MUNSON, W.A. (1933)
"Loudness, its definition, measurement and calculation", J.A.S.A., vol. 5, pp. 82-108.
- GREEN, D.M. (1973)
"Minimum integration time", Basic Mechanisms in Hearing, edited by Moller, A.R., Academic Press New-York, pp. 829-846.
- HANSEN, J.A. (1972)
"Redresseurs efficaces", Revue technique Bruel et Kjaer, N° 2, pp. 3-18.
- KATZMAN, F.L. (1976)
"How to choose and use an RMS meter correctly", Electronics, March 18, pp. 97-100.

- KIM, J. et Mc NEILAGE, P. (1972)
"Voiceless consonant durations VCV utterances", 84th Meeting of the A.S.A., Miami, Nov. 1972, 5 p.
- KILATT, D.H. (1976)
"Linguistic uses of segmental duration in english : Acoustic and perceptual evidence", J.A.S.A., vol. 59, pp. 1208-1221.
- LEHISTE, H. (1970)
Suprasegmentals, Chapter 2, MIT Press, pp. 6-53.
- LIENARD, P. (1978)
Decibels et indices de bruits, Masson, Paris, p. 53.
- ROSSI, M. (1971)
"L'intensité spécifique des voyelles", Phonetica, vol. 24, pp. 129-161.
- ROSSI, M. (1976)
"La perception des variations d'intensité", Travaux de l'Institut de Phonétique d'Aix, vol. 3, pp. 361-457.
- SMATHERS, T. (1971)
"A unique monolithic AGC/ squelch amplifier", Linear Application I, An 51, National Semiconductor Corporation, Santa Clara.
- STEVENS, S.S. (1961)
"Procedure for calculating loudness function", J.A.S.A., vol. 33, pp. 1577-1585.
- TESTON, B. (1971)
"Description d'un système de détection de l'intensité et de la mélodie de la voix", Travaux de l'Institut de Phonétique d'Aix, vol. 1, pp. 129-145.
- TESTON, B. et ROSSI, M. (1977)
"Un système de détection automatique du fondamental et de l'Intensité", Actes des 8èmes Journées d'étude sur la parole, G.A.L.F., Aix, pp. 111-117.
- VAN ERCK, M.H. et RAUCH, S. (1976)
"How to measure a.c signal accurately", Electronics, March 18, pp. 94-96.
- VAN ERCK, M.H. (1977)
"Mesures de valeurs alternatives et efficaces", T et M News Philips Supplément I, p. 12.
- WARHMAN, C.G. (1958)
"A true RMS instrument", Technical review Bruel et Kjaer, N° 3, pp. 9-21.

- WAHRMAN, C.G. et BROCH, J.T. (1975)
"Sur le temps d'intégration des mesures en valeur efficace",
Revue technique Bruel et Kjaer, N° 2, pp. 3-23, N° 3, pp. 3-36.
- WAHRMAN C.G. (1978)
"RMS and LMS circuit developments", Bruel et Kjaer Technical review
-à paraître-.
- WONG, Y.J. et OTT, W.E. (1976)
Function Circuits Chapter 4, Mac Graw Hill, New-York, pp. 126-176.
- ZWICKER, E. (1960)
"Ein verfahren zur bestimmung der lautstärke", Acustica, vol. 10, pp.304.
- ZWICKER, E. (1973)
"Temporal effects in Psychoacoustical excitation", Basic Mechanisms in Hearing, edited by Moller, A.R. Academic Press New-york, pp. 829-846.
- ZWISLOCKI, J.J. (1969)
"Temporal summation of loudness : an analysis", J.A.S.A., vol. 46,
pp. 431-441.