



ELSEVIER
MASSON

Disponible en ligne sur
ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com

IRBM

IRBM xxx (2010) xxx-xxx

Article original

Approche empirique pour déterminer les générateurs neurophysiologiques sous-jacents des potentiels évoqués auditifs engendrés par des sons de parole (*Speech ABR*)

Empirical approach to find the neurophysiological generators of speech auditory brainstem response

I. Akhoun^{a,b,c}, C. Berger-Vachon^{b,c,d,e,*}, L. Collet^{b,c,d}

^a *Audiology and Deafness Group, Human Communication and Deafness, School of Psychology, Faculty of Medicine and Human Sciences, Ellen Wilkinson Building, University of Manchester, Manchester, M13 9PL, Royaume-Uni*

^b *CNRS-UMR 5020 « neurosciences sensorielles, comportement et cognition », équipe audiology, université de Lyon, université Claude-Bernard Lyon-1, 69003 Lyon, France*

^c *IFR 19, institut fédératif des neurosciences de Lyon, 69677 Bron, France*

^d *Pavillon U, service audiology et exploration orofaciale, hôpital Edouard-Herriot Lyon-3, 69003 Lyon, France*

^e *Filière génie-biomédical, institut des sciences et techniques de l'ingénieur de Lyon, université de Lyon, université Claude-Bernard Lyon-1, bâtiment Omega, avenue du 11 Novembre 1918, 69100 Villeurbanne, France*

Reçu le 31 octobre 2008 ; accepté le 15 décembre 2009

Résumé

Les potentiels évoqués auditifs en réponse à un son de parole (*Speech Auditory Brainstem Response [Speech ABR]*) sont une méthode de mesure objective non-invasive de l'activité neurophysiologique auditive au niveau du tronc cérébral. Les générateurs neurophysiologiques de cette réponse sont méconnus, mais les latences, ainsi que les caractéristiques spectrales du *Speech ABR* suggèrent que ceux-ci font partie du tronc cérébral supérieur (entre le noyau cochléaire et le colliculus inférieur). Au vu des caractéristiques fonctionnelles bien particulières des unités cellulaires des voies auditives, différents stimuli acoustiques ont été choisis, afin de mettre en évidence la sensibilité des *Speech ABR* aux caractéristiques acoustique de ces stimuli, et par suite, d'en déduire les probables sites neurophysiologiques pouvant avoir généré cette réponse. Ainsi, nous avons enregistré des *Speech ABR* en réponse à des sons purs, des sons complexes harmoniques, des syllabes /ba/, /pa/, ainsi que leurs homologues (somme de sinusoides à la fréquence des formants modulées par l'enveloppe temporelle des syllabes). Par ailleurs, le modèle *Auditory Image Model* (Patterson et al., 1995 [17]), qui simule l'activité neurale en entrée du colliculus inférieur, montre que les homologues comme les syllabes donnent lieu à une quantité d'activité comparable contrairement aux FFR, suggérant ainsi que le processus neurophysiologique engendrant la FFR se situe au-delà de la périphérie cochléaire. En effet, l'apparente synthèse des informations de toute la gamme fréquentielle acoustique aboutissant à une FFR oscillatoire et synchronisée sur la période fondamentale rappelle les caractéristiques fonctionnelles des cellules *disc-shape* du *colliculus inférieur central*, telles qu'elles ont été décrites dans le modèle de codage de la périodicité (Voutas et al., 2005 [42]).

© 2010 Publié par Elsevier Masson SAS.

Mots clés : Audition ; Electrophysiologie ; Parole ; Potentiels évoqués auditifs ; *Frequency following response*

Abstract

Auditory evoked potentials to speech (*Speech auditory brainstem response [Speech ABR]*) are a non-invasive way to investigate neurophysiological activity, at the level of the brainstem. The *Speech ABR* precise neurophysiological generators remain poorly defined. However, latencies and low-pass spectrum both suggest that these generators might lie in the upper brainstem (roughly between the cochlear nucleus and the inferior colliculus). Having considered the particular functional pattern of cells along the auditory pathway, specific stimuli have been synthesized to make out the acoustic sensitivity of *Speech ABR* components. Accordingly, hypotheses have been made on the probable neurophysiological generators, most likely to have elicited both *Speech ABR* components: onset response and frequency following response. *Speech ABR* have been recorded to

* Auteur correspondant.

Adresses e-mail : Idrick.Akhoun@manchester.ac.uk (I. Akhoun), cbv@univ-lyon1.fr (C. Berger-Vachon).

pure tones, harmonic complex tones, /ba/ and /pa/ syllables, and their analogues (calculated as a sum of five weighted sine waves at the formant frequencies and amplitudes, and modulated by the syllables temporal envelopes). In addition, the Auditory Image Model (Patterson et al., 1995 [17]), simulating the neural activity at the auditory periphery, i.e. inferior colliculus input, suggests that both analogues and syllables elicit the same amount of energy, in contrast to the recorded FFR. This contradiction means that the neurophysiological signal processing leading to FFR is made beyond auditory periphery. Indeed, FFR synchronisation on F0 seems to be the result of an overall processing of the whole stimulus spectrum. This behaviour reminds the functional characteristics of disc-shape cells in the inferior colliculus, as described in a previous study of physiological periodicity coding (Periodicity analysis network, Voutsas et al., 2005 [42]).

© 2010 Published by Elsevier Masson SAS.

Keywords: Audition; Electrophysiology; Speech; Auditory brainstem response; Frequency following response

Glossaire :

ABR (Speech ABR) : potentiel évoqué auditif provenant du tronc cérébral. Le *Speech ABR* est engendré par un stimulus de parole ; dans le cas d'une syllabe consonne-voyelle, il se compose d'une réponse *onset* suivie d'une réponse en fréquence (*frequency following response* [FFR]).

Homologue : Son non-signifiant synthétisé pour ressembler aux syllabes dans les basses fréquences (même enveloppe), avec une structure fine très dégradée. Il est important de se rappeler que ces homologues n'avaient pas pour but de ressembler perceptivement aux syllabes (vu l'appauvrissement spectral appliqué au-delà de 500 Hz et leur inharmonicité).

Bark : échelle de fréquences reproduisant la répartition de la sensibilité fréquentielle sélective de la membrane basilaire (tonotopie cochléaire, voire *code neural spatial*). Cette échelle est linéaire dans les fréquences en dessous de 1 kHz, logarithmique au-delà.

Pour passer de l'échelle linéaire (Hz) à l'échelle Bark :

$$Z(\text{Bark}) = 6,7 * \arcsinh((f(\text{Hz}) - 20) / 600)$$

Réciproquement :

$$f(\text{Hz}) = \sinh(Z(\text{Bark}) / 6,7) * 600 + 20$$

Code neural temporel (code neural spatial) : le code neural temporel est lié au *phase-locking*. Il utilise la synchronisation des neurones pour véhiculer des informations sur la périodicité du stimulus acoustique. Le code neural spatial découle de la tonotopie cochléaire (décomposition fréquentielle du plus aigu au plus grave des sons le long de la membrane basilaire, comme sur le clavier d'un piano) ; le long des voies auditives, les neurones ont souvent une fréquence caractéristique à laquelle ils déchargent le mieux, et ainsi, le son décomposé en fréquence dans la cochlée est véhiculé dans des canaux (filtres) auditifs passe-bandes. Notons que ces deux codes sont complémentaires, comme le montre la FFR dans cette étude (reflet du code neural temporel ne fonctionnant que dans des conditions précises de code spatial – afférences des différents filtres auditifs).

Enveloppe : contour d'un stimulus, estimé dans cette étude par la méthode valeur absolue – filtrage passe bas (500 Hz choisi ici par rapport à la fréquence de coupure du *speech ABR*). Notons que l'enveloppe peut également être estimée par transformée de Hilbert (Gilbert et Lorenzi, 2006 [47]).

Frequency following response : partie pseudopériodique de l'ABR, engendrée par une portion pseudopériodique du stimulus acoustique. Reflet de l'activité de *phase-locking composite*.

Grande moyenne : les ABR sont moyennés un grand nombre de fois (environ un millier) par sujet. La grande moyenne est la moyenne de cette moyenne, pour tous les sujets.

Hauteur tonale : sensation auditive liée à la fréquence (de grave à aiguë).

Insert earphones : équivalent du casque, mais le transducteur acoustique (transformant le signal audio en onde acoustique) est encastré dans des boîtiers, reliés à l'oreille du sujet par des tubes en caoutchouc et des *earplugs*. L'avantage est d'emprisonner les artefacts électromagnétiques [17]. Mais les tubes d'air peuvent distordre le signal (de façon marginale toutefois avec des tubes de 30 cm).

Mesures objectives : estimation de la perception du sujet, sans avoir à solliciter sa description perceptive (psycho-acoustique).

Réponse de latence moyenne : réponse d'origine thalamocorticale intervenant environ entre 20 et 80 ms après l'indice du stimulus qui lui a donné naissance.

Réponse onset : partie transitoire de l'ABR, engendrée par une impulsion acoustique. Couramment utilisé en routine clinique (dites ondes I, III, V).

Phase-locking : propriété de certains neurones de synchroniser leurs décharges de potentiels d'action sur la périodicité du stimulus acoustique.

Prémoyennage : la fenêtre d'acquisition commence l'enregistrement avant le début du stimulus, pendant la période de prémoyennage. Cette portion de l'ABR, où aucune réponse n'est attendue, permet de se faire une idée du bruit de recueil ambiant.

Structure fine acoustique : contrairement à l'enveloppe, il s'agit des composantes spectrales hautes-fréquence (habituellement au-delà de 500 Hz). Elle peut également être déduite de la transformée de Hilbert.

1. Introduction

1.1. Contexte

Les potentiels évoqués auditifs (PEA) permettent d'enregistrer l'activité électrique résultant de la propagation et du traitement du message nerveux auditif dans les noyaux successifs des voies auditives [1]. L'enregistrement des PEA est très simple à mettre en œuvre et implique une chaîne de stimulation auditive et une chaîne de recueil électrophysiologique provenant de trois électrodes de contact. Le signal issu des PEA dépend à la fois de la structure acoustique du stimulus, de la durée de l'acquisition ainsi que de la gamme fréquentielle des filtres d'acquisition et enfin (ce qui en fait son intérêt clinique), de l'état fonctionnel des voies auditives du sujet. L'exemple le plus simple est celui d'un normo-entendant stimulé par une

Pour citer cet article : Akhoun I, et al. Approche empirique pour déterminer les générateurs neurophysiologiques sous-jacents des potentiels évoqués auditifs engendrés par des sons de parole (*Speech ABR*). IRBM (2010), doi:10.1016/j.irbm.2009.12.005

impulsion acoustique (clic). Le signal recueilli consiste en une série de pics successifs [2] qui traduisent la réponse des noyaux successifs des voies auditives (un ms : ganglion spiral; trois ms : noyau cochléaire; cinq ms : lemme latéral, colliculus inférieur, corps genouillé médian dans le thalamus, au-delà de 100 ms : cortex auditif).

L'absence, ou la détérioration d'un des pics du PEA peut donc être reliée à un dysfonctionnement du traitement du signal auditif à un endroit donné des voies auditives [3]. Le personnel clinique en audiologie a couramment recours aux PEA (en appoint de la description du patient) pour le diagnostic de la surdité, ou pour évaluer l'efficacité des techniques de réhabilitations auditives comme par exemple les implants cochléaires [4,5]. Ces PEA sont très utiles pour avoir un aperçu de l'activité auditive transitoire, en réponse à l'impulsion; mais un son de parole du type consonne-voyelle peut considérablement élargir l'horizon d'étude [6,7].

La syllabe /ba/, par exemple, possède deux indices acoustiques remarquables : une impulsion acoustique (transition consonne-voyelle) suivie d'une onde acoustique pseudopériodique (la voyelle). Ces deux segments vont solliciter deux parties distinctes des voies auditives [8,9]. En effet, anatomiquement et fonctionnellement, les noyaux du tronc cérébral sont répartis en différents types de cellules, dont le code neural temporel est soit sensible aux transitions acoustiques, soit démontre une faculté à synchroniser la décharge de potentiels d'action sur la périodicité du stimulus (phénomène dit de calage de phase ou *phase-locking*) [10]. Or, ces deux aspects du code neural temporel ont un impact important sur l'acuité perceptive et dans l'encodage de la parole [11]. En pratique, les PEA en réponse à la parole (*Speech Auditory Brainstem Response [Speech ABR]*) sont un signal électrophysiologique enregistré sur le scalp, résultant de la somme de toutes ces activités sous-jacentes. Le *Speech ABR* a été décrit comme une succession de trois composantes [7,9] :

- le pré-moyennage, enregistré pendant une période de silence précédant le stimulus et permettant d'avoir une idée du niveau de bruit de recueil ambiant ;
- la réponse transitoire (réponse *onset*), sensible aux transitions acoustiques (consonne-voyelle, par exemple). D'une façon générale, ses caractéristiques sont proches de l'onde V des PEA engendrés par un clic ;
- la réponse en fréquence (*frequency following response [FFR]*) qui reproduit fidèlement l'enveloppe du stimulus acoustique et traduit le phénomène de *phase-locking* sous-jacent [8,9,12,13].

1.2. Problématique

Le *Speech ABR* permet donc d'avoir un double aperçu de l'activité transitoire (réponse *onset*) et synchronisée sous-jacente (FFR), d'où son potentiel intérêt pour évaluer la qualité du traitement auditif temporel dans le tronc cérébral. Toutefois, il est fondamental de maîtriser deux aspects de cette technique, dans le but de la valider et d'en permettre l'utilisation

clinique, qui fait pour l'instant défaut. Premièrement, il est nécessaire de pouvoir corréler la sensibilité de ces mesures électrophysiologiques et la sensibilité perceptive (psychophysique) des sujets [14]. Deuxièmement, il est crucial de cibler les sites neurophysiologiques d'où provient la réponse. Or, si cette démarche d'investigation in situ est répandue chez l'animal anesthésié, elle est très lourde à implémenter chez l'humain (éveillé).

L'exploration du traitement de l'information auditive dépend de l'échelle d'étude : depuis l'histogramme post-stimulus jusqu'à l'analyse psycho-acoustique de la perception sonore qui engage le système auditif dans son ensemble. D'un côté, une description anatomofonctionnelle précise du système auditif a pu être extraite des études chez l'animal [10] et de l'autre, la psychoacoustique a permis d'obtenir une connaissance détaillée de l'acuité auditive chez le normo-entendant comme chez le malentendant [15]. De même, la neuro-imagerie corticale permet une approche de l'intégration du signal auditif [16]. Toutefois, d'une part, les données de l'animal, a fortiori anesthésié, ne prennent pas en compte des interactions probables du type excitation-inhibition, et d'autre part, la neuro-imagerie ne permet pas d'observer précisément le fonctionnement des voies auditives subcorticales. Par conséquent, il existe donc un angle mort dans la fenêtre d'observation des mécanismes neurophysiologiques subcorticaux en jeu chez l'humain éveillé. Or la connaissance précise de ces mécanismes subcorticaux serait d'autant plus intéressante que le signal auditif transite obligatoirement dans le tronc cérébral avant d'atteindre le cortex et d'être consciemment perçu, en subissant un prétraitement important. Le *Speech ABR* se présente comme un moyen de combler ce manque de connaissance du traitement auditif subcortical.

Dans cette étude, nous avons utilisé différents stimuli et sélectivement manipulé leurs indices acoustiques afin de mettre en évidence des analogies avec les caractéristiques fonctionnelles de certaines cellules du système auditif. Nous avons enregistré des *Speech ABR* en réponse à des sons purs, des sons complexes harmoniques, des syllabes /ba/ /pa/, ainsi que leurs homologues (somme de sinusoides à la fréquence des formants modulées par l'enveloppe temporelle des syllabes). La FFR fut absente pour les sons purs ou les homologues, alors qu'une FFR très nette fut enregistrée pour les sons complexes harmoniques ou pour les syllabes. Ce résultat paradoxal montre que la FFR, bien que prenant la forme de l'enveloppe temporelle du stimulus, est influencée par toute la gamme fréquentielle du stimulus, en particulier par les hautes fréquences. D'autre part, le modèle *auditory image model* [17], qui simule l'activité neurale en entrée du colliculus inférieur (générateur supposé de la FFR), montre que les homologues comme les syllabes donnent lieu à une quantité d'activité comparable contrairement aux FFR, suggérant ainsi que le processus neurophysiologique engendrant la FFR se situe au-delà de la périphérie cochléaire. En effet, l'apparente synthèse des informations de toute la gamme fréquentielle acoustique aboutissant à une FFR oscillatoire et synchronisée sur la période fondamentale rappelle les caractéristiques fonctionnelles des cellules *disc-shape* du colliculus inférieur central, telles qu'elles ont été décrites dans le modèle de codage de la périodicité [42].

2. Matériel et méthodes

2.1. Sujets

Les ABR ont été enregistrés sur dix sujets pour la première étude et 15 pour la seconde. Tous avaient une audition normale (seuils auditifs inférieurs à 20 dB HL). Ces expérimentations se sont déroulées en accord avec la loi Huriet et elles furent approuvées par le comité d'éthique de Lyon (CCPPRB # 05-024). Les sujets ont notamment signé un consentement de participation.

2.2. Stimuli acoustiques

Dans l'expérience 1, les stimuli utilisés furent;

- six sons purs sinusoïdaux ;
- deux sons complexes harmoniques ayant pour fréquence fondamentale (F_0) respectivement 110 et 220 Hz et composé chacun de dix harmoniques non déphasés.

Chaque son durait 100 ms et était pondéré quasi-linéairement, à l'aide du logiciel libre Audacity[®], en ouverture et en fermeture sur 10 ms. La fréquence des sons purs (choisis pour balayer la gamme fréquentielle la plus importante pour la perception de la parole) était: 130, 220, 440, 880, 1760 et 3520 Hz.

Dans l'expérience 2, les stimuli étaient:

- deux syllabes enregistrées /ba/ et /pa/ situées aux deux extrémités d'un continuum de *voice onset time* (VOT) [18];
- leurs homologues.

Le VOT est l'écart de temps séparant l'ouverture du conduit vocal du début du voisement. Ces deux syllabes avaient la même voyelle et ne différaient l'une de l'autre que par la durée de la consonne et la zone de coarticulation (VOT respectivement de -55 et 10 ms). La durée des phonèmes était 230 ms, et la voyelle, identique dans les deux cas, commençait à 140 ms. Dix millisecondes de pré-moyennage précédaient le début du phonème /ba/ et cent quarante pour le /pa/, afin de quantifier le bruit ambiant. Le fondamental (F_0) était stationnaire autour de 200 Hz et les quatre formants (F_1 à F_4) variaient quant à eux autour de 750, 1660, 2800 et 4400 Hz.

Les homologues du /ba/ et du /pa/ furent déduits de ces caractéristiques acoustiques pour avoir la même enveloppe temporelle, tout en ayant une structure fine appauvrie et très simplifiée. Ces homologues furent synthétisés sous Matlab (the Mathworks[®]) comme la somme de cinq sinusoïdes (non déphasées) modulées par l'enveloppe temporelle de la syllabe correspondante. Ces cinq sinusoïdes avaient respectivement pour fréquence: F_0 , F_1 , F_2 , F_3 et F_4 (soit le fondamental et les quatre formants de la voyelle /a/). De la même façon, ces sinusoïdes furent pondérées par a_0 , a_1 , a_2 , a_3 et a_4 , soit l'amplitude spectrale correspondante de la voyelle /a/. Enfin, cette somme de sinusoïdes fut modulée par l'enveloppe temporelle de la syllabe correspondante ($Env(t)$), obtenue d'après la valeur absolue du signal de parole, filtrée passe-bas à 500 Hz par un filtre à cor-

rection de phase de Butterworth d'ordre 4). La formule de ces homologues est donc de la forme :

$$\text{Homologue } (t) = Env(t)/a_0 \cdot [a_0 \sin(2\pi F_0 t) + a_1 \sin(2\pi F_1 t) + a_2 \sin(2\pi F_2 t) + a_3 \sin(2\pi F_3 t) + a_4 \sin(2\pi F_4 t)]$$

Avec:

$$F_0 = 200, F_1 = 750, F_2 = 1660, F_3 = 2800, F_4 = 4400 \text{ [Hz]}$$

$a_0 = 751$, $a_1 = 193$, $a_2 = 33$, $a_3 = 4,3$, $a_4 = 4,3$ (unités arbitraires)

Notons que l'homologue ainsi synthétisé ne ressemble perceptivement pas à la syllabe de départ (ce qui apparaît peu surprenant au vu de l'appauvrissement spectral au-delà de 500 Hz); gardons à l'esprit que ces homologues furent synthétisés uniquement dans le but d'avoir la même structure temporelle (i.e. la même enveloppe) et d'en déduire son influence sur l'ABR engendré.

2.3. Mesure des ABR

Les ABR ont été enregistrés sur le Centor USB (Racia-Alvar[®]) par des électrodes d'argent, comme c'est, par ailleurs, le cas en routine clinique dans le service d'audiologie pour l'acquisition des clic-ABR. L'électrode active fut placée sur le vertex (sommet du crâne) et les électrodes de référence sur la mastoïde droite (voie 1) et gauche (voie 2). L'électrode de masse fut placée sur le front. Lors des expérimentations, la valeur des impédances fut maintenue en dessous de 5 k Ω .

Le dispositif expérimental utilisé fut optimisé pour éliminer toute influence des artefacts électromagnétiques (Fig. 1), comme nous l'avons précédemment décrit [19]. En voici le détail.

2.3.1. Chaîne de stimulation

Les stimuli utilisés furent joués par le logiciel Audacity[®] par blocs de 1200 traces dans l'expérience 1 (800 traces dans l'expérience 2), à une cadence de 6,6 stimuli par seconde (respectivement 3). La stimulation était binaurale à un niveau de 45 dB *Sensation Level* (au-dessus du seuil du sujet), ce qui représente un niveau sonore confortable, voire fort, en fonction des sujets.

Le niveau sonore était maîtrisé par un audiomètre AC-30 (Interacoustics[®]), qui reliait la carte-son Soundblaster Audigy[®] aux transducteurs acoustiques: des inserts-earphones (Biologic[®]). Les inserts étaient blindés par une cage de Faraday reliée à la masse électrique du dispositif expérimental afin d'éliminer le risque d'artefact électromagnétique; le son étant diffusé au bout d'une paire de tubes de caoutchouc. Notons que le transfert de la stimulation acoustique dans les tubes d'air long de 30 cm induit 1 ms de retard, qui fut ensuite retranché dans les enregistrements.

2.3.2. Chaîne d'acquisition

Pour l'expérience 1, chaque ABR fut enregistré dans une fenêtre d'analyse de 150 ms (300 ms pour l'expérience 2), composée de 2000 points (respectivement 4000); d'où une fréquence d'échantillonnage de 13,3 kilohertz. Les ABR ainsi enregistrés étaient filtrés en temps réel entre 80 et 3200 Hz, afin de se concentrer sur la gamme d'activité des potentiels du tronc cérébral auditif. Chaque trace trop ample (au-delà de

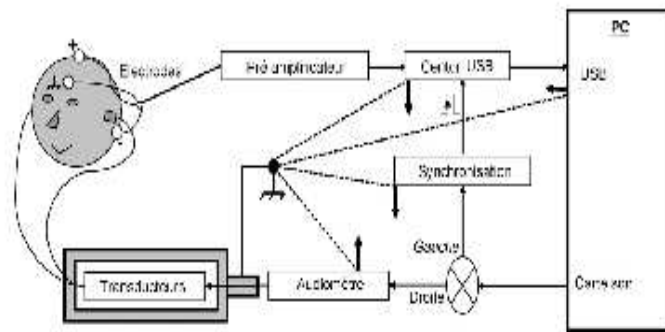


Fig. 1. Dispositif expérimental utilisé pour l'enregistrement des ABR. Il se compose d'une chaîne de recueil (haut) et d'une chaîne de stimulation (bas). Les deux chaînes sont synchronisées par une impulsion (dite trigger) située sur la voie de gauche du fichier son et reliée à l'entrée de synchronisation de l'appareil d'enregistrement des ABR Centor USB, (adapté d'après [19]).

75 μ V) était assimilée à un artefact musculaire et de ce fait rejetée.

Dans le but d'éliminer tout artefact lié au potentiel microphonique cochléaire [20], nous avons enregistré les ABR engendrés par polarité alternée, c'est-à-dire le stimulus et son opposé. La transduction sensorielle dans l'organe de Corti qui donne naissance au signal auditif est très peu sensible à la polarité, tandis que les artefacts susceptibles de parasiter l'ABR reproduisent le stimulus tel quel en suivant la polarité. Par conséquent, la somme des ABR dans chaque polarité (dite polarité alternée) réduit les artefacts sans affecter la réponse neurophysiologique auditive. Dans l'expérience 2, un ABR en polarité alternée est en fait le moyennage de 1600 traces (800 traces de chaque polarité).

2.4. Analyse a posteriori des ABR

Les ABR furent analysés visuellement d'après leurs traces temporelles, puis leur contenu spectrotemporel fut analysé en temps-fréquence.

Les grandes moyennes temporelles combinent les ABR de tous les sujets. Elles furent calculées comme la somme algébrique des données individuelles dans une polarité donnée (positive, négative, alternée). Auparavant, les ABR furent normalisés (soustraction de la composante continue et correction des erreurs standard par loi normale centrée).

Les spectrogrammes (d'énergie) furent obtenus sous Matlab (par transformée de Fourier à court terme sur des fenêtres glissantes de 1024 points pour les ABR et 2048 pour les stimuli acoustiques, se recouvrant à 90%), à partir des traces temporelles normalisées. Par conséquent, 298 spectres furent obtenus pour les ABR sur 513 bandes de fréquences larges de 13 Hz ; de manière homologue, pour les stimuli, 223 spectres sur 2049 bandes de fréquences de 11 Hz. Les grandes-moyennes des spectrogrammes furent ensuite obtenues en faisant la moyenne des spectrogrammes individuels. Les composantes spectrotemporelles furent considérées présentes si l'énergie dans cette région était supérieure de 6 dB à celles environnantes. Des sections temporelles (coupes horizontales) de ces spectrogrammes permettaient de mesurer la durée et l'amplitude de la FFR, sur

la bande de fréquences correspondant à la fréquence de synchronisation, c'est-à-dire F_0 . Notons que les informations de phase ont été volontairement ignorées afin d'éviter d'atténuer artificiellement les hautes fréquences de la FFR lors du grand-moyennage. En effet, nous avons constaté que des déphasages significatifs (dits jitter) peuvent exister entre FFR pour différents sujets. Dans le cas où ce jitter est suffisamment important pour que les FFR soient en opposition de phase, il se peut que ces FFR disparaissent finalement après grand-moyennage. Cela fut le cas en particulier pour les stimuli de fréquences élevées [21]. Enfin, l'activité globale dans les bandes de fréquences du spectrogramme a été estimée en sommant algébriquement les énergies contenues dans ces bandes (ce qui donne un nombre sans dimension pertinente [dB. ms]).

2.5. Description des modèles auditifs

Si les enregistrements in situ sont très compliqués à mettre en place pour situer les générateurs neurophysiologiques des *Speech ABR* chez l'humain, les latences ainsi que l'aspect passé-bas de la FFR suggèrent que ceux-ci sont susceptibles de se situer dans le tronc cérébral supérieur [7]. Par conséquent, nous avons utilisé le modèle *Auditory image model* (AIM) [17] afin de simuler la réponse électrophysiologique du système auditif périphérique (entre la cochlée et le noyau cochléaire) dont les projections se situent principalement dans le tronc cérébral supérieur. Ce modèle permet donc ainsi de simuler le signal en entrée des (supposés) générateurs sous-jacents de la FFR. Notons cependant qu'il est probable que d'importantes interactions efférentes aient lieu chez l'humain éveillé entre le cortex auditif et le tronc cérébral supérieur. Cette interaction encore méconnue n'est pas prise en compte dans ce modèle, issu de données animales. En voici le principe de fonctionnement.

Le modèle AIM (version 2006) est composé d'une série de modules en cascade qui simulent les différentes étapes du traitement de l'information auditive. Notons que cette simulation basée sur des études animale in situ ne peut se borner qu'à donner une idée des mécanismes en jeu. Cette série de modules est organisée de la façon suivante [22]:

Pour citer cet article : Akhoun I, et al. Approche empirique pour déterminer les générateurs neurophysiologiques sous-jacents des potentiels évoqués auditifs engendrés par des sons de parole (*Speech ABR*). IRBM (2010), doi:10.1016/j.irbm.2009.12.005

- le stimulus acoustique (fichier wave) subit tout d'abord un traitement précochléaire (*precochlear processing* [PCP]) qui reproduit les étapes de filtrage passe-bande entre 500 et 15000 Hz du conduit auditif externe (pavillon de l'oreille externe et chaîne d'osselets dans l'oreille moyenne) [23];
- dans le module suivant (*basilar membrane motion* [BMM]), l'étrier entraîne la vibration de la membrane basilaire, où le stimulus est décomposé en fréquence (hautes fréquences à la base de la cochlée et basses fréquences à l'apex). Cette décomposition spectrale opère sous la forme d'un banc de 50 filtres auditifs gammatones se recouvrant et répartis entre 100 et 6000 Hz [17]. Ces filtres ont une réponse impulsionnelle de la forme : $g(t) = a t^{b-1} \exp(-2\pi b \text{ERB}(f_c) t) \cdot \cos(2\pi f_c t + \Phi)$, cela étant valable pour $t > 0$. Les $\text{ERB}(f_c)$ représentent la largeur de bande des filtres rectangulaires (équivalents à l'action des filtres auditifs mesurés psycho-acoustiquement à un niveau sonore modéré), soit : $\text{ERB}(f_c) = 24,7 + 0,108 f_c$ en Hz [24]. La distance entre ces filtres auditifs est estimée en *Bark*. Notons que lors de l'encodage cochléaire, les filtres auditifs centrés sur les hautes-fréquences sont stimulés avant ceux des basses-fréquences, en raison du retard dû à la propagation de l'onde mécanique le long de la membrane basilaire de la base vers l'apex;
- lors de cette décomposition fréquentielle, les cellules ciliées internes encodent l'amplitude de la vibration mécanique cochléaire sous la forme d'un train de potentiels d'action se propageant sur le nerf auditif dans chaque canal (filtre) auditif. C'est le troisième module (*neural activity pattern* [NAP]), qui retrace la naissance du message nerveux. Une particularité de cette transduction est que les cellules ciliées redressent le signal filtré dans le module PCP (les décharges de potentiels d'action n'ont lieu que dans un sens du mouvement de la membrane basilaire) et lui appliquent un filtrage passe-bas, vu leur temps de réponse limité pour les vibrations trop rapides en raison de la nécessité de renouveler les transmetteurs synaptiques [25,26]. Le module NAP tient également compte des phénomènes d'adaptation et de réfraction des nerfs, qui ont tendance à moins décharger quand ils sont trop sollicités. De plus, il est tenu compte de la disparition du *phase-locking* dans les hautes-fréquences (au-delà de 5 kHz) dans ce module. Notons enfin le parallèle entre ce mécanisme et la façon dont nous avons calculé l'envolpe de nos stimuli dans l'expérience 2, ainsi que le rationnel de l'utilisation de la polarité alternée. Le NAP est considéré comme une modélisation de l'activité neurale en entrée des générateurs neurophysiologiques du *Speech ABR*;
- le modèle propose ensuite de modéliser l'action du tronc cérébral supérieur et du cortex auditif, à travers les images auditives stabilisées (*stabilised auditory images* [SAI]). Ces centres nerveux intégrés sont censés assimiler l'information véhiculée dans tous les filtres auditifs, afin d'en extraire les informations les plus pertinentes pour comprendre le son. Par exemple, une auto-corrélation de l'activité neurale conglomerée de tones les sorties des filtres auditifs du module NAP donne des maxima dont l'écart est relié à la période de hauteur tonale perçue par le sujet [26,27]. Toutefois, nous ne nous y intéressons pas, car c'est précisément cette activité

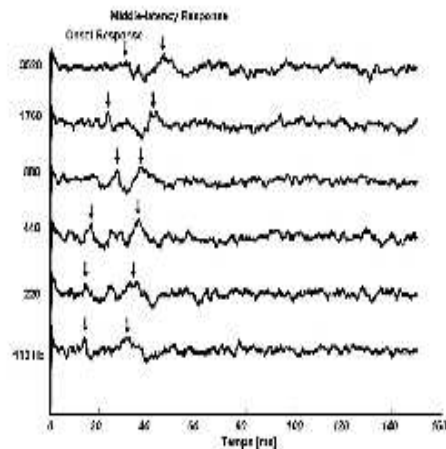


Fig. 2. Grande moyenne temporelle ($n = 10$) des ABR en réponse à des sons purs. La fréquence des sons purs ayant engendrée ces ABR est indiquée à gauche. La déflexion au début est due à un artefact du système d'acquisition. La seule réponse identifiable est l'onde positive de latence moyenne située entre 35 (à 110 Hz) et 55 ms (à 3520 Hz). Elle ne correspond ni à une réponse *onset*, ni à une FFR. Cette dernière est clairement absente. En revanche, les pics situés entre 15 (à 110 Hz) et 25 ms (à 3520 Hz) s'apparentent à une réponse *onset*.

que nous pensons enregistrer à travers les *Speech ABR*. Autrement dit, nous considérons la sortie du module NAP comme une simulation de l'entrée de notre « système » (générateurs neurophysiologiques des *Speech ABR*).

3. Résultats

3.1. ABR en réponse à des sons purs et sons complexes harmoniques

Quels que soient les sons purs utilisés, aucune réponse caractéristique du *Speech ABR* ne fut obtenue, tandis que les deux sons complexes harmoniques ont donné lieu à une FFR très nette.

Il est frappant de constater qu'absolument aucune FFR ne fût enregistrée pour les sons purs, comme l'atteste la platitude de la courbe Fig. 2, et cela, quelle que soit la fréquence du stimulus. En effet, l'amplitude spectrale à F_0 et ses trois premiers harmoniques ne fut pas significativement différente du bruit de recueil¹; test t apparié, $t = -0,83$, $p = 0,4$ pour 130 Hz, $t = -0,45$, $p = 0,6$ pour 220 Hz. En revanche, quelques pics suggèrent la présence d'une réponse *onset*, entre 12 ms (pour 110 Hz) et 22 ms (à 3520 Hz). Mais sachant que le stimulus commençait après 10 ms de pré-moyennage, ces réponses *onset* (attendues une dizaine de millisecondes après la transition acoustique du début du stimulus) 110 et 220 Hz sont peu vraisemblables. Toutefois, l'onde la plus remarquable dure une dizaine de millisecondes et se situe entre 30 ms (à 110 Hz) et 45 ms (à 3520 Hz). Elle correspond à une réponse d'origine thalamocorticale dite de latence moyenne, [28]. Il est intéressant de noter que la latence de cette réponse de latence moyenne

¹ Le bruit ambiant de contrôle par rapport à F_0 et ses trois harmoniques fut estimé par l'amplitude spectrale à $1,5 F_0$, $2,5 F_0$, $3,5 F_0$ et $4,5 F_0$.

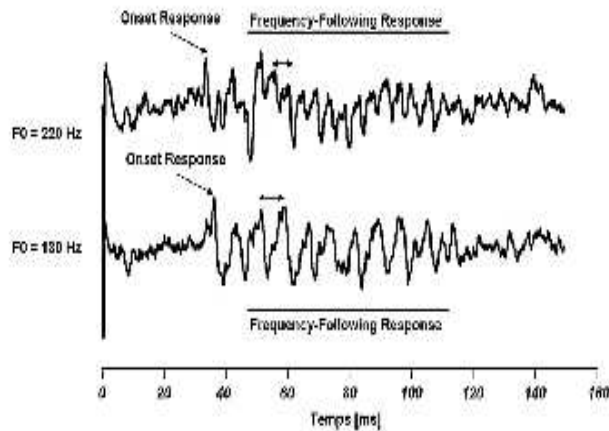


Fig. 3. Grande moyenne temporelle ($n=10$) des ABR en réponse à des sons complexes harmoniques. À l'inverse de la figure précédente (sons purs), la FFR est très clairement présente ici. Elle est synchronisée sur F_0 (précisée par les doubles-flèches). La réponse *onset* est bien présente au début du stimulus. On ne remarque pas de réponse de latence moyenne ici.

semble augmenter graduellement avec la fréquence des stimuli.

Dans le cas des sons complexes harmoniques, l'ABR présente une forme classique, avec une réponse *onset* suivie d'une FFR. La FFR fut périodique, bien synchronisée sur F_0 (Fig. 3). L'analyse spectrotemporelle (non-reportée ici), montra que les harmoniques de F_0 furent présents jusqu'à une fréquence de coupure entre 400 et 600 Hz. À l'inverse des sons purs, l'amplitude spectrale à F_0 et ses trois premiers harmoniques fut significativement différente du bruit de recueil: test t apparié, $t=6,88$, $p<0,001$, pour le son complexe harmonique à 130 Hz de fondamental, $t=6,08$, $p<0,001$, pour celui à 220 Hz.

3.2. ABR en réponse à des syllabes /ba/ et /pa/, et à leurs homologues

Ici, nous n'avons observé aucun effet de la polarité et de la latéralité². Par conséquent, toutes les données seront présentées en polarité alternée et sur la voie 1. Le résultat phare est que les ABR obtenus en réponse à /ba/, comme à /pa/ ont montré des FFR bien définies, contrairement aux homologues (Fig. 4).

3.2.1. Réponses onset

Plus précisément, les homologues ont donné lieu à des réponses *onset*, situées peu après l'apparition de l'équivalent

² En effet, les coefficients d'intercorrélation entre les grandes moyennes pour les deux polarités fut maximal pour des décalages de zéro ms pour tous les stimuli ($R=0,27, 0,13$ et $0,37$ pour /ba/, homologue /ba/ et l'homologue /pa/). Notons que pour la condition /pa/, le coefficient de corrélation maximal ($R=0,12$) fut mesuré pour un décalage de 4,8 ms, mais ce pic était proche d'un autre centré en zéro ($R=0,11$). De plus, les ABR ainsi obtenus furent similaires sur les deux canaux d'enregistrements (droite, voie 1 et gauche, voie 2 en fonction de la position des électrodes). Cela reste visible sur les traces temporelles (Fig. 4) et cette conjecture est confirmée par intercorrélation entre les ABR de chaque voie, qui fut maximal pour un décalage de zéro ms ($R=0,47, 0,43, 0,36, 0,63$ pour /ba/, /pa/, homologue /ba/ et homologue /pa/). Cela suggère une absence de latéralisation visible, tout du moins avec la technique des ABR.

de la voyelle, entre 130 et 150 ms. À l'inverse, peu de réponses *onset* furent visibles en réponse à /ba/ (le pic situé vers 50 ms ressemble à une réponse *onset*, mais s'apparente davantage à une réponse de latence moyenne, comme dans le cas des sons purs), alors qu'un pic existe à 170 ms pour le /pa/. Cette réponse *onset* est également visible sur les spectrogrammes et se présente comme un rectangle foncé (donc énergétique) dans les basses fréquences entre 100 et 150 ms, et entre 50 et 200 Hz (Fig. 5). L'analyse statistique de l'énergie contenue dans ce rectangle foncé montre que cette région spectrotemporelle est plus énergétique qu'une région contrôle située avant la voyelle pour les syllabes, comme pour les homologues (entre 50 et 100 ms, et entre 50 et 200 Hz): test t apparié, $t(/ba/)=6,1$, $t(/pa/)=4,0$, $t(\text{homologue } /ba/)=2,7$, $t(\text{homologue } /pa/)=5,1$, $p<0,05$ dans tous les cas (Fig. 6).

3.2.2. Différentes FFR pour les syllabes et leurs homologues

Comme prévu, la FFR (lorsqu'elle fut enregistrée pour /ba/ ou /pa/, mais pas pour leurs homologues) fut concentrée dans les basses fréquences, avec une forte composante à F_0 et à ses harmoniques jusqu'à environ 600 Hz. Ces composantes spectrales sont visibles sur les spectrogrammes de la FFR par les barres horizontales blanches dans les bandes de fréquences correspondantes, pendant la durée de la voyelle. Toutefois, dans la condition /pa/, des composantes de FFR furent obtenues jusqu'à 1000 Hz, allant ainsi au-delà des fréquences de coupure habituellement reportées [7,9,13]. Notons enfin que les composantes fortement énergétiques à 900 ou 1200 Hz sont le résultat de l'influence distordue de l'artéfact électromagnétique lié à l'alimentation électrique (50 Hz).

Les FFR en réponse à /ba/ et à /pa/ furent semblables. Cette conjecture visuelle (Fig. 4) a été démontrée par un test t non significatif, $p(/ba/ \text{ vs } /pa/)=0,35$, entre l'énergie de la FFR contenue dans la bande spectrale de F_0 , que ce soit pour /ba/ (qui s'échelonnait parmi les 15 sujets entre -30 et -9 dB, $M=-18$ dB, $std=6$ dB), ou pour /pa/ (entre -31 et -9 dB,

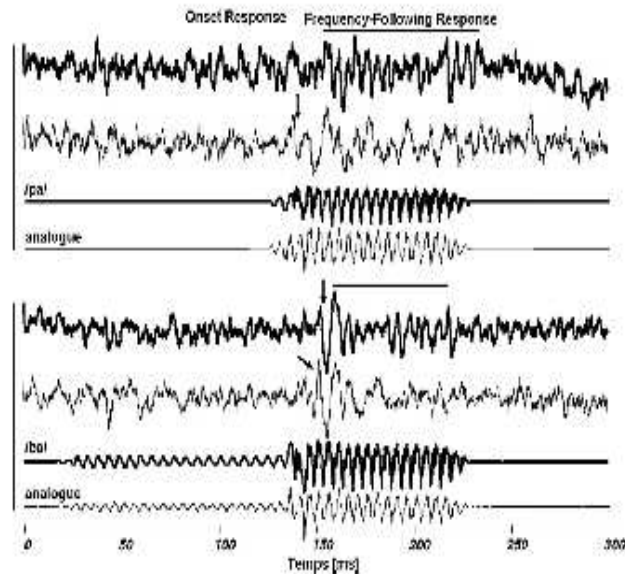


Fig. 4. Comparaison (temporelle) des speech- et homologues-ABR. La syllabe /pa/ est représentée en gras, ainsi que la grande moyenne ($n=15$) des Speech ABR. L'homologue du /pa/ (structure temporelle identique, mais très fortement appauvrie spectralement) est représenté de la même façon en trait fin. De manière homologue se situe en dessous tout ce qui correspond au /ba/.

$M = -17$ dB, $std = 6$ dB). En revanche, la différence entre la FFR obtenue pour les syllabes par rapport à leurs homologues fut très marquée, entre /ba/ et l'homologue /ba/, dont l'énergie dans la bande de F_0 variait entre -37 et -17 dB ($M = -17$ dB, $std = 5$ dB), et pour l'homologue /pa/, entre -33 et -20 dB ($M = -25$ dB, $std = 3$ dB). En effet, chaque syllabe et son homologue eurent des amplitudes spectrales à F_0 significativement différentes, $p(/pa/ \text{ vs } \text{homologue})$ inférieur à 0,001 et $p(/ba/ \text{ vs } \text{homologue})$ égal à 0,04.

Pour détailler cette différence, chaque syllabe et leur homologue ont été comparées (sur certaines bandes de fréquences) afin de mieux évaluer les conséquences du contenu spectral en entrée des générateurs de la FFR sur le contenu spectral résultant sur la FFR. Dans un premier temps, nous avons représenté la distribution d'énergie dans les bandes de fréquences (i.e. coupes-horizontales du spectrogramme) du spectrogramme les plus pertinentes quant au stimulus utilisé : $F_0 = 200$ Hz et ses deux harmoniques (Fig. 7). Il est bien visible sur la partie gauche de la Fig. 7 que les syllabes (en gras) /ba/, comme /pa/ et leurs homologues (en traits fins) ont la même quantité d'énergie à F_0 pendant la portion voisée du son, mais que les syllabes sont plus énergétiques pour les fréquences supérieures. À l'inverse, sur la partie droite de la Fig. 7, on remarque que la FFR était bien plus énergétique dans la bande de fréquences de F_0 (200 Hz) lorsqu'elle est engendrée par les syllabes /ba/, comme /pa/ (traits gras), par rapport aux homologues (traits fins) dans la zone de la voyelle (rectangles en pointillés). Pour les fréquences supérieures, les syllabes comme les homologues présentent des répartitions énergétiques comparables.

Dans un deuxième temps, la différence entre les syllabes et leurs homologues fut comparée pour quatre critères sur un plus grand nombre de fréquences sur les diagrammes de la Fig. 8a

(et idem sur la Fig. 8b). Sur ces deux figures, les différences d'énergie dans huit bandes de fréquences données (100 [bruit de recueil], 200 (F_0), 400, 600, 800, 1000, 1200 et 1400 Hz) ont été mesurées pour les quatre critères suivants. Tout d'abord, nous avons mesuré :

- la différence d'énergie présente dans ces coupes de spectrogrammes entre les grandes moyennes des ABR obtenues en réponse au /ba/ (resp /pa/) et à son homologue ;
- puis de la même façon, sur les ABR individuels ;
- ensuite, sur le même principe, nous avons calculé la différence d'énergie présente sur ces coupes de spectrogrammes pour les stimuli acoustiques du /pa/ et de son homologue.

Enfin, nous avons calculé la différence d'activité neurale (modélisée par AIM, sur le module NAP) engendrée par /pa/ et par son homologue. Le résultat principal de cette comparaison de contenu spectral est que, conformément à ce qui est décrit Fig. 7, on remarque un contraste à 200 Hz (F_0) entre le comportement des ABR et les caractéristiques spectrales des stimuli acoustiques ou bien de l'activité neurale modélisée par AIM :

- d'un côté, les ABR sont significativement plus énergétiques pour /ba/ par rapport à son homologue (Anova: $F(1,118) = 4,811$, $p = 0,03$) et même de façon encore plus franche pour le /pa/ (Anova: $F(1,118) = 16,6$, $p < 0,001$) ; et la grande moyenne, qui accentue davantage la FFR en améliorant le rapport signal sur bruit par rapport aux données individuelles, accentue ce résultat ;
- d'un autre côté, les stimuli acoustiques /ba/, comme /pa/, sont similaires pour F_0 (ils furent synthétisés de fait pour avoir la même enveloppe, d'où cette périodicité identique) ;

Pour citer cet article : Akhoun I, et al. Approche empirique pour déterminer les générateurs neurophysiologiques sous-jacents des potentiels évoqués auditifs engendrés par des sons de parole (Speech ABR). IRBM (2010), doi:10.1016/j.irbm.2009.12.005

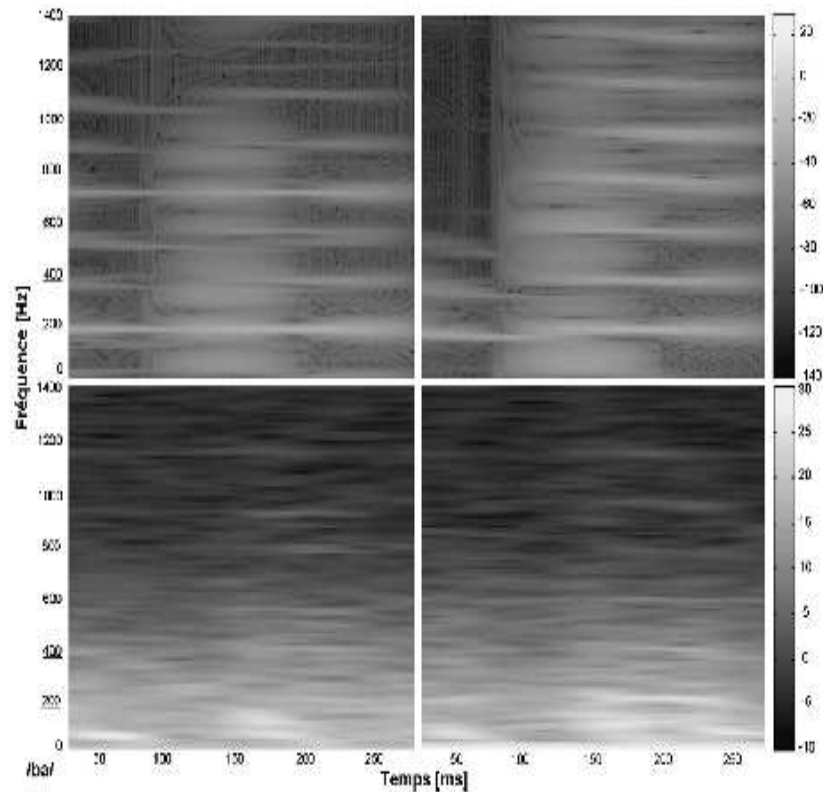


Fig. 5. Comparaison spectrotemporelle des speed- et homologues-ABR, pour le /ba/. Représentation temps-fréquence du vrai /ba/ (haut-droite) et de son homologue (haut-gauche), ainsi que des grandes moyennes des spectrogrammes des ABR correspondants (en dessous). Le vrai /ba/ et son homologue ont la même enveloppe temporelle, ce qui se voit ici dans la similarité entre leurs composantes à la fréquence du fondamental, tandis que la structure fine (harmoniques, formants) est beaucoup plus dynamique et progressive pour le vrai /ba/. Le /ba/ a donné lieu à une FFR bien définie (bande fréquentielle blanche à 200 Hz entre 150 et 250 ms).

- il est intéressant de constater que le modèle neural AIM (module NAP) simule en effet une activité neurale différente pour les syllabes et leurs homologues (Fig. S1–13). Ce modèle prédit une activité quasiment semblable pour le /ba/, comme pour le /pa/ et leurs homologues aux alentours de F_0 , tandis que plus de disparité est attendue pour les fréquences plus élevées. Notons toutefois qu'il ne s'agit pas de code neural temporel mais de code spatial, à l'inverse de la FFR : les fréquences mentionnées ne représentent pas (nécessairement) une activité oscillatoire des neurones impliqués, mais elles représentent la fréquence centrale des filtres auditifs.

4. Discussion

Une meilleure connaissance des origines anatomofonctionnelles des ABR serait d'une grande utilité en vue d'implémenter cet outil dans le diagnostic des pathologies auditives, mais c'est également un des rares moyens techniques à disposition pour enregistrer l'activité auditive subcorticale chez l'humain éveillé. Le résultat-clé de cette étude est que deux sons ayant la même périodicité n'engendrent pas systématiquement de FFR. Ainsi, nous n'avons pas obtenu de FFR avec des sons purs, tandis que des sons complexes harmoniques ont engendré des

FFR très franches. Et de plus, nous avons obtenu des FFR très franches en réponse à des syllabes /ba/ et /pa/, mais pas de FFR en réponse à leurs homologues (malgré le fait que leurs enveloppes soient identiques). Ce distinguo entre FFR et caractéristiques acoustiques, qui n'apparaît pas dans la modélisation de l'activité neurale en entrée des générateurs supposés de la FFR, fait pencher la balance vers une hypothèse de codage global de toutes les composantes harmoniques sous la forme d'une synchronisation sur F_0 . Ce type de fonctionnement suggère un rapprochement avec les unités fonctionnelles détectrices de coïncidences, correspondant anatomiquement aux cellules *disc-shape* du colliculus inférieur central [42].

4.1. Origine de la réponse onset et son rôle dans le traitement de la parole

La réponse *onset* est semblable aux clics-ABR, qui sont bien connus et couramment utilisés en routine clinique d'exploration fonctionnelle des voies auditives subcorticales [2]. La réponse *onset* semble bien être le résultat d'une réponse transitoire liée à une transition acoustique qui s'apparente à un clic (telle que le passage entre l'équivalent de la consonne à celui de la voyelle dans le cas de l'homologue /ba/). L'origine neu-

Pour citer cet article : Akhoun I, et al. Approche empirique pour déterminer les générateurs neurophysiologiques sous-jacents des potentiels évoqués auditifs engendrés par des sons de parole (Speech ABR). IRBM (2010), doi:10.1016/j.irbm.2009.12.005

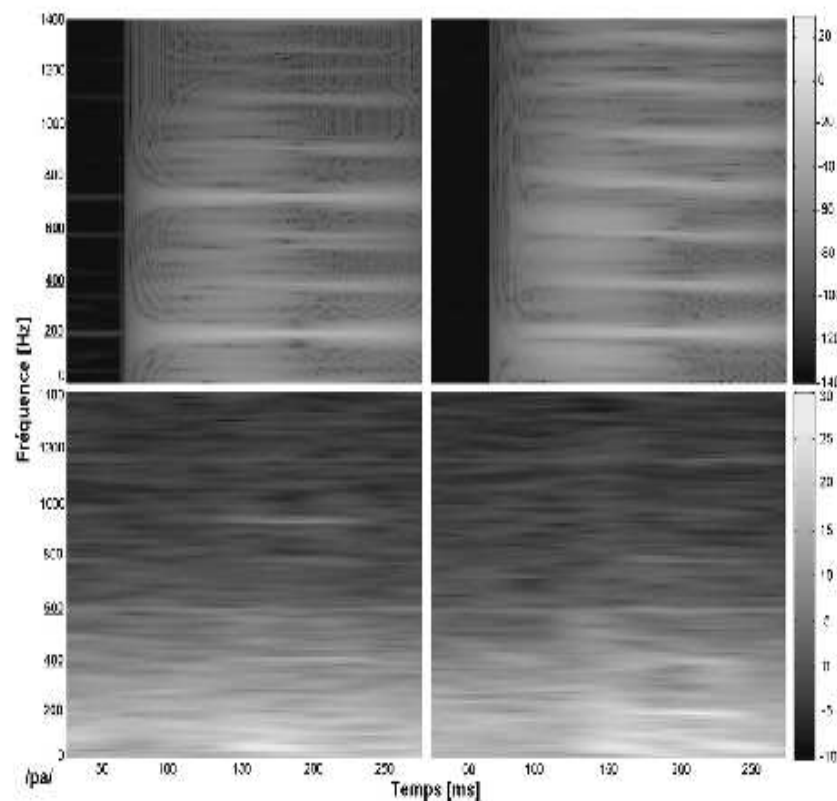


Fig. 6. (idem Fig. 5) Comparaison spectrotemporelle des speech- et homologues-ABR, pour le /pa/. La similarité des voyelles entre le /ba/ et le /pa/ est bien visible en comparant les fenêtres en haut à droite sur cette figure et la précédente. On pourrait ajouter qu'on ne voit pas particulièrement d'effet de la consonne en comparant les Fig. 5 et 6.

rophysiologique de l'onset response semble plus triviale que celle de la FFR : il s'agirait de l'expression d'un grand nombre de décharges neurales simultanées, consécutif au passage d'un segment acoustique de faible intensité (consonne) à un autre segment beaucoup plus énergétique (voyelle). Fonctionnellement, la majorité des unités cellulaires du tronc cérébral possèdent des caractéristiques de décharge (décrites dans [10,11,42]) qui sont susceptibles d'intervenir dans la naissance de la réponse onset. On peut citer entre autres les cellules oscillatrices (cellules en étoile du noyau cochléaire postéroventral : fonction *chopper*), les cellules interrupteur (cellules pieuvres ou multipolaires du noyau cochléaire postéroventral : fonction *pure onset* ou *onset chopper*), voire les détecteurs de coïncidence (cellules disques du colliculus inférieur central, fonction : soutenue [*sustained*]), en définitive la plupart sont concernées, sauf les cellules intégratrices (cellules fusiformes du noyau cochléaire dorsal, fonction : *pauser buildup*). Notons au passage qu'une partie des cellules ayant participé à cette décharge massive de potentiels d'action se retrouvera en état réfractaire pendant une période de quelques millisecondes, ce qui explique en partie pourquoi la réponse onset des ABR est plus importante que la FFR.

Du fait de sa sensibilité à la transition consonne-voyelle, la réponse onset a parfois constitué un espoir de pouvoir servir de mesure objective du VOT. Le VOT est un indice acoustique

primordial dans la perception de la parole pour discriminer par exemple le /ba/ (VOT = -55 ms) du /pa/ (20 ms pour la langue française). En revanche, dans cette étude, nous n'avons pas été à même d'observer de différences visibles entre les réponses onset engendrées par le /ba/ ou le /pa/, contrairement aux homologues. De plus amples investigations, mettant par exemple en jeu un continuum entre deux paires de phonèmes, permettraient d'en savoir plus sur le lien entre la réponse onset et le VOT, que nos données ne permettent pas d'entrevoir clairement.

4.2. Non-réciprocité de la relation entre la FFR et l'enveloppe temporelle du stimulus : importance des hautes-fréquences

Que ce soit dans les études animales in situ ou dans les études électrophysiologiques chez l'humain éveillé, la FFR a été décrite comme une onde synchronisée sur la périodicité du stimulus. D'un point de vue physiologique, la FFR est un potentiel évoqué illustrant l'activité globale de *phase-locking* dans les cellules du tronc cérébral, entre le noyau cochléaire et le corps genouillé médian du thalamus [29-37]. Notons que la spectaculaire synchronisation entre la FFR et la périodicité du stimulus nécessite de s'assurer de l'authenticité de l'origine neurale de cette réponse, afin de prouver qu'il ne s'agit ni d'artéfact

Pour citer cet article : Akhoun I, et al. Approche empirique pour déterminer les générateurs neurophysiologiques sous-jacents des potentiels évoqués auditifs engendrés par des sons de parole (Speech ABR). IRBM (2010), doi:10.1016/j.irbm.2009.12.005

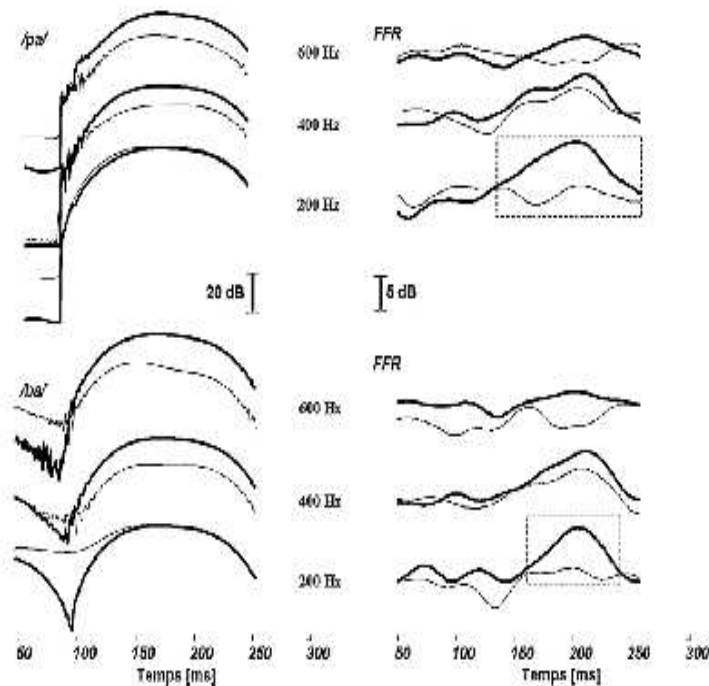


Fig. 7. Comparaison des coupes des spectrogrammes ci-dessus (Fig. 5 et 6). Comparaison de l'énergie présente dans les bandes de fréquences correspondant au fondamental (200 Hz) et les deux premiers harmoniques (400 et 600 Hz). Les signaux concernant la syllabe sont en gris et en trait fin pour l'homologue. Les stimuli se situent à gauche et la FFR à droite. Enfin, ce schéma concernant le /ba/, en haut, est repris à l'identique pour le /pa/ en bas. On note que la similarité des enveloppes des homologues et des syllabes se retrouve dans la superposition entre les courbes d'énergie des stimuli à 200 Hz (que ce soit pour le /ba/ ou pour le /pa/). En revanche, la distribution d'énergie est différente pour les harmoniques supérieurs. À l'intérieur, l'énergie contenue dans la FFR à 200 Hz est beaucoup plus important en réponse aux syllabes (comme souligné par les rectangles en pointillés); tandis que celle-ci est semblable dans les deux conditions pour les harmoniques 400 et 600 Hz.

électromagnétique, ni de potentiel microphonique cochléaire [3, 19]. Dans une récente étude, nous avons rapporté que la FFR imitait très fidèlement l'enveloppe temporelle du stimulus /ba/ (redressée-filtrée à 500 Hz) et que le spectre fréquentiel de la FFR ainsi que celui de la voyelle étaient identiques jusqu'à une fréquence de coupure de 500 Hz [9]. Peut-on pour autant dire que, réciproquement, seule l'enveloppe temporelle est impliquée dans la FFR? Cette hypothèse a été élégamment réfutée par l'étude de Greenberg et al. [33]. Dans leurs travaux, la FFR a été décrite comme une mesure objective de la sensation de hauteur tonale perçue, notamment dans le cas du phénomène psychoacoustique de fondamentale manquante³ (*missing fundamental*). Le système auditif est donc capable de reconstituer l'information liée à F_0 à partir des hautes fréquences. Pour ce faire, deux hypothèses sont envisageables liées au code spatial ou temporel :

- soit toutes les informations disponibles dans chacun des filtres auditifs sont combinées pour résulter en un message lié à F_0 (la hauteur tonale perçue) ;

- soit la forme des fluctuations de l'enveloppe temporelle⁴ incarne l'information de hauteur tonale.

Toutefois, cette dernière hypothèse « temporelle » a été mise en cause et il a été montré que le code spatial a plus d'impact que le code temporel dans la perception de la hauteur tonale [38]. La question se pose alors de connaître la fonctionnalité du phénomène auditif qui résulte en une FFR.

Dans la continuité, les résultats de nos deux expériences apportent un éclaircissement sur ce que signifie la synchronisation de la FFR. L'absence de FFR pour les homologues et la FFR typique que nous avons mesurées pour les syllabes /ba/ (ou idem pour /pa/), alors que ces stimuli ont la même enveloppe temporelle montre d'une autre façon (par rapport à [33]) que la FFR ne se borne pas à simplement véhiculer les informations acoustiques de l'enveloppe temporelle. De plus, la FFR typique obtenue en réponse aux sons complexes harmoniques (avec F_0 à 110 ou 220 Hz) et l'absence de FFR pour chacune des harmoniques isolés est révélatrice. Il semblerait que la FFR soit sensible à la totalité du spectre fréquentiel et qu'elle traduise un

³ Dans ce cas, un son complexe harmonique dont toutes les harmoniques sont perçues, mais que F_0 a été retirée, est toujours perçue à une hauteur tonale correspondant à F_0 (quand bien même cette composante fréquentielle est physiquement absente).

⁴ Les fluctuations de l'activité neuronale se font à une périodicité de F_0 lorsque des battements apparaissent du fait que deux harmoniques successifs soient codés dans le même filtre auditif, on les dit alors *non-résolus*.

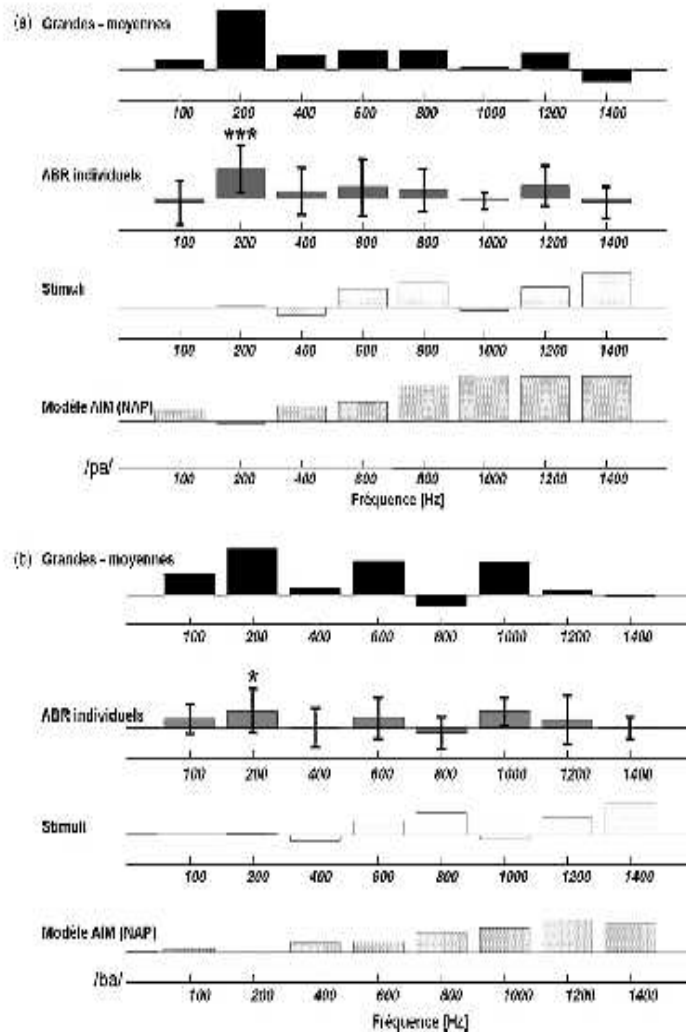


Fig. 8. a : comparaison globale des composantes fréquentielles entre la /pa/ et son homologue pour quatre situations : (i) grande moyenne des FFR ; (ii) FFR individuelles ; (iii) stimuli acoustiques ; (iv) modèle AIM (module NAP). Ces valeurs furent mesurées aux huit fréquences 100 (en dessous de F_0), 200 (F_0), 400, 600, 800, 1000, 1200 et 1400 Hz. Chaque barre dans ces histogrammes représente la différence entre la syllabe et son homologue pour les énergies sommées de la bande de fréquences correspondante du spectrogramme pour les quatre situations (i) à (iv) ci-dessus. On remarque que la contradiction de la composante 200 Hz (F_0), encadrée en pointillés est très marquée pour les (i) et (ii) (la FFR était donc plus énergétique à 200 Hz en réponse à la syllabe par rapport à son homologue), contrairement à (iii) et (iv) (la modélisation NAP qui représente l'entrée supposée du système neurophysiologique de synthèse des ABR, comme les stimuli, montrent davantage d'activité dans les hautes fréquences pour la syllabe par rapport à l'homologue, alors que l'activité en réponse aux deux est similaire dans les basses fréquences) ; b : idem Fig. 8a pour /ba/.

phénomène auditif dans lequel la totalité des harmoniques est considérée dans leur ensemble et non pas individuellement.

Il semblerait donc que la FFR, si elle existe, soit de la forme d'une onde synchronisée sur la périodicité du stimulus ; mais elle est conditionnée par la totalité du contenu fréquentiel du stimulus (et par extension, de la totalité des informations véhiculées dans chaque filtre auditif) et en particulier la structure fine (contenu acoustique fréquentiel dans les fréquences au delà de 500 Hz environ). Cela recoupe les observations de Dau [39], qui, dans une modélisation de la FFR chez l'humain conçue à partir d'un modèle de décharge du nerf auditif, a montré que les filtres auditifs hautes-fréquences (au delà de 1,5 kHz) étaient les

plus importants dans la formation de la FFR. En effet, l'onde acoustique sur la cochlée (où les basses fréquences sont codées spatialement à l'apex) subissent une dispersion dans le processus de propagation qui les déphase et rendent le message nerveux en résultant moins bien synchronisé.

D'après la modélisation NAP d'AIM (présentée en annexe), la quantité d'information neurale qui transite dans chaque bande de fréquences semble ne pas être distribuée de la même façon pour la voyelle du vrai son de parole par rapport à l'homologue : pour le vrai son de parole, chaque bande de fréquences possède des maxima d'activité neurale pour des écarts entre deux pics d'activité (abscisses) progressif, tandis que ces écarts sont

Pour citer cet article : Akhoun I, et al. Approche empirique pour déterminer les générateurs neurophysiologiques sous-jacents des potentiels évoqués auditifs engendrés par des sons de parole (Speech ABR). IRBM (2010), doi:10.1016/j.irbm.2009.12.005

les mêmes sur la plupart des bandes de fréquences dans le cas de l'homologue. De façon imagée, on peut dire que l'activité neurale en « croses de hockey » telles qu'on peut les observer pour le vrai son de parole est plus à même de donner lieu à une forte synchronisation de la FFR que l'activité neurale avec des écarts fixes entre les pics d'activité. Les générateurs de la FFR sembleraient jouer le rôle d'intégrateurs de cette activité périphérique (les afférences doivent arriver de façon régulière plutôt que tout d'un coup) et qu'ils agiraient de préférence comme les détecteurs de coïncidences⁵ (nécessité d'avoir des afférences de différents filtres auditifs). Il fut, par ailleurs, intéressant de constater qu'une étude récemment publiée spécule également sur ce mécanisme physiologique, en se référant aux travaux de Langner [40].

4.3. Paradoxe de la FFR, qui intuitivement, ne devrait pas pouvoir être enregistrée au niveau du scalp

Le rôle supposé des détecteurs de coïncidences dans la FFR apporte un indice quant à la question de savoir comment est-ce que la FFR parvient de façon aussi claire sur les électrodes placées sur le scalp, malgré la « cacophonie neurale » ambiante qui accompagne le trajet du message nerveux auditif, au moins pour les raisons suivantes :

- tous les cellules du système auditif sont spécialisées et ont notamment des caractéristiques fonctionnelles spécifiques telles que les taux de décharge spontanés, des fréquences caractéristiques de réponse, ainsi que des types de décharge particuliers (transitoires, périodiques synchronisées comme les oscillatrices, périodiques mais avec une périodicité intrinsèque comme les détecteurs de coïncidences, etc...);
- dès le nerf auditif, chaque filtre auditif (illustré ici par les 50 signaux du modèle AIM) répond de façon décalée en raison du délai de propagation de la vibration de la membrane basilaire. Or intuitivement, on s'attendrait à ce que cette activité ne puisse être enregistrée en champ lointain (sur le scalp), seulement à condition que ces flux afférents du signal auditif, dans ces canaux parallèles, soient synchronisés et non pas déphasés;
- ce déphasage a aussi lieu dans la « dimension perpendiculaire » à la précédente. En effet, chaque noyau cellulaire (du nerf auditif au colliculus inférieur et au delà) répond également au fur et à mesure, avec des latences progressives d'une milliseconde en sortie de la cochlée jusqu'à une dizaine de millisecondes dans le colliculus inférieur.

⁵ Il se pourrait que dans le cas des homologues, le fait que la plupart des bouffées d'activités neuronales aient lieu pour des mêmes intervalles de temps doit entraîner une sur-stimulation des cellules du noyau cochléaire (ce que modélise le module NAP d'AIM) impliquées dans ce codage (codeur d'enveloppe et intégrateurs), qui pourrait les rendre réfractaires; dans ce cas, les détecteurs de coïncidences (générateurs suspects des Speech ABR) recevraient moins d'afférences, ce qui entraînerait une réponse moindre, et in fine, une FFR très amoindrie enregistrée au niveau du scalp.

Néanmoins, Eggermont [1] a récemment décrit les phénomènes électriques qui font que le message nerveux peut malgré tout parvenir jusqu'au scalp dans le cas des clic-ABR (ce qui peut être vraisemblablement étendu à la FFR). La dynamique de transmission du message et l'exposition topographique des générateurs neurophysiologiques sont particulièrement importants. En d'autres termes, afin d'être correctement recueilli, le message nerveux doit profiter des changements de conductivité du milieu biologique traversé (tissus/os) et dépend beaucoup de l'alignement des cellules. C'est au passage pourquoi l'auteur invalide l'impact du thalamus pour donner des réponses enregistrables sur le scalp, car les cellules auditives sont masquées par l'activité ambiante des autres types de cellules environnantes.

Les raisons de la netteté de la FFR que nous enregistrons peut-être étayée par cette hypothèse de la nécessaire dynamique de la source électrophysiologique. Elle pourrait découler de la singularité du code neural en jeu dans les cellules du colliculus inférieur (dit *rate modulation transfert function* [rMTF]): son principe est de coder l'information de modulation de fréquence par un taux de décharge proportionnel à la fréquence modulante [10,41]. À l'inverse, dans les noyaux cellulaires situés en amont des voies auditives entre la cochlée et le colliculus inférieur, le type de codage n'est que du type *temporal modulation transfer function* (tMTF): certaines cellules déchargent chacune selon leurs caractéristiques spécifiques et leur taux de décharge est synchronisé sur celui du stimulus. En somme, avec le rMTF, les cellules ont un taux de décharge variable, qui varie avec une périodicité synchronisée sur F_0 tandis que pour le tMTF, les cellules ont un taux de décharge constant à F_0 . Intuitivement, on pourrait penser qu'un phénomène dynamique tel que le rMTF, qui n'apparaît qu'à partir du colliculus inférieur serait privilégié pour résulter en une différence de potentiel enregistrable sur le scalp, ce qui réglerait le problème du déphasage des réponses entre les différents noyaux du système auditif. Par ailleurs, pour résoudre le problème du déphasage des filtres auditifs, on peut raisonnablement supposer que les cellules du colliculus inférieur, en particulier les détecteurs de coïncidence, combinent les informations de plusieurs filtres auditifs, par ailleurs certainement en prenant en compte cette information de phase.

5. Conclusion : implication clinique

L'étude du comportement de la FFR en fonction des caractéristiques spectrales et temporelles des stimuli utilisés a permis de cibler les générateurs potentiels de la FFR, en se référant aux particularités fonctionnelles des cellules du tronc cérébral, décrites à partir des données animales. L'exacte connaissance des générateurs neurophysiologiques de la FFR pourrait permettre de faire le lien entre une forme anormale des *Speech ABR* et un dysfonctionnement de cette partie du colliculus inférieur. Cela serait en particulier très utile pour parfaire le diagnostic des surdités centrales ou des neuropathies auditives [43].

En plus de cibler un groupe cellulaire, ces deux expériences ont montré l'influence majeure de la structure fine acoustique sur la FFR : la FFR produit une oscillation qui imite la forme de l'enveloppe du stimulus, à partir de la synthèse de l'ensemble des informations spectrales. Or, de récentes études ont mis en avant

le rôle capital de la structure fine acoustique dans la perception de la parole [44,45] et en particulier chez les malentendants [46]. De plus, les auteurs ont montré que les informations d'enveloppe qui sont primordiales dans la perception de la parole peuvent être extraites de la structure fine [47]. De plus amples investigations seraient toutefois nécessaires pour évaluer la possibilité pour le *Speech ABR* d'être une mesure objective de la perception de la parole, ou plus précisément, de la correcte utilisation par le patient mal-entendant des indices acoustiques contenus dans la structure fine.

Conflit d'intérêt

Aucun conflit d'intérêt n'est déclaré ici.

Remerciements

L'expérience 2 de ce travail a été présentée en poster [48] au congrès « Acoustics 2008 », à Paris.

Chaleureux remerciements aux sujets ayant participé à cette étude, ainsi qu'aux collègues pour leur point de vue précieux : Colette McKay et Chris Plack (Manchester), Thomas Bridonneau, Clément Barral, Frédéric Marmel, A & K Nyamay'Antu, Barbara Tillmann, Nicolas Grimault (Lyon), Arnaud Norena, Yves Cazals (Marseille), Étienne Gaudrain et Tim Ives (Cambridge), pour l'aide autour de l'utilisation d'AIM. Un grand merci à Céline Richard (CHU Bellevue, Saint-Étienne) pour son aide dans la mise en place et l'exécution des enregistrements. Annie Moulin, Hung Thai-Van et Evelyne Vuillet ont également beaucoup contribué à ce travail. Les regards de Jean-Luc Schwartz (Grenoble) et Paul Avan (Clermont-Ferrand), en tant que rapporteurs de thèse d'où cette étude est extraite, ont été autant passionnants qu'infiniment utiles. La relecture minutieuse d'un lecteur anonyme a permis d'améliorer beaucoup une version antérieure de ce manuscrit. Idrick Akhoun a été financé par un contrat doctoral CIFRE entre Cochlear France, le CNRS et l'UCBL, en étant co-encadré par Christian Berger-Vachon et Hung Thai-Van. Il est maintenant en poste à Manchester, où il est impliqué dans un projet de recherche sur les potentiels évoqués composites du nerf auditif, sous la direction de Colette McKay.

Annexe A. Matériel complémentaire

Le matériel complémentaire accompagnant la version en ligne de cet article est disponible sur doi:10.1016/j.irbm.2009.12.005.

Références

- [1] Eggermont JJ. Electric and magnetic fields of synchronous neural activity: Peripheral and central origins of AEP. Auditory evoked potentials. Basic principles and clinical applications. In: Burkard R, Don M, Eggermont, JJ (Eds.), (Philadelphia, Baltimore, New-York, London, Buenos Aires, Hong Kong, Sydney & Tokyo. Lippincott Williams & Wilkins) 2007.
- [2] Picton T, Hillyard S, Krausz H, Galambos R. Human auditory evoked potentials. I. Evaluation of components. *Electroenceph clin Neurophysiol* 1974;36(2):179–90.

- [3] Marsh JT, Worden FG, Smith JC. Auditory Frequency following response: neural or artifact? *Science* 1970;169(3951):1222–3.
- [4] Gallego S, Frachet B, Michéyl C, et al. Cochlear implant performance and electrically-evoked brain-stem response characteristics. *Electroencephalogr Clin Neurophysiol* 1998;108:521–5.
- [5] Brown CJ. Clinical uses of electrically evoked auditory nerve and brainstem responses. *Curr Opin Otolaryngol Head Neck Surg* 2003;11:383–7.
- [6] Kraus N, Nicol T. Brainstem origins for cortical what and where pathways in the auditory system. *Trends Neurosci* 2005;28(4):176–81.
- [7] Johnson KL, Nicol TG, Kraus N. Brainstem response to speech: a biological marker of auditory processing. *Ear Hear* 2005;26(5):424–34.
- [8] Krishnan A. Frequency following response. Auditory evoked potentials. Basic principles and clinical applications. In: Burkard R, Don M, Eggermont, JJ (Eds.), (Philadelphia, Baltimore, New-York, London, Buenos Aires, Hong Kong, Sydney & Tokyo. Lippincott Williams & Wilkins) 2007.
- [9] Akhoun I, Gallégo S, Moulin A, Vuillet E, Berger-Vachon C, et al. The temporal relationship between Speech auditory brainstem responses and the acoustic pattern of the phoneme /ba/. *Clin Neurophysiol* 2008;119:922–33.
- [10] Joris PX, Schreiner CE, Rees A. Neural processing of amplitude modulated sounds. *Physiol Rev* 2004;84(2):541–77.
- [11] Palmer AR, Shamma SA. Physiological representations of speech. *Springer Handbook of Auditory Research*, (vol. 18) Speech processing in the auditory system. In: Greenberg S, Ainsworth WA, Popper AN, Fay RR (Eds.), (New-York. Springer) 2004.
- [12] Galbraith GC. Two-channel brain-stem frequency-following responses to pure tone and missing fundamental stimuli. *Electroencephalogr Clin Neurophysiol* 1994;92(4):321–30.
- [13] Wible B, Nicol T, Kraus N. Atypical brainstem representation of onset and formant structure of speech sounds in children with language-based learning problems. *Biol Psychol* 2004;67:299–317.
- [14] Purcell DW, John SM, Schneider BA, Picton TW. Human temporal auditory acuity as assessed by envelope following responses. *J Acoust Soc Am* 2004;116(6):3581–93.
- [15] Moore BC. An Introduction to the psychology of hearing, 5th ed. (Bradford. Emerald Group Publishing) 2004.
- [16] Moore BC. Basic auditory processes involved in the analysis of speech sounds. *Philos Trans R Soc Lond B Biol Sci* 2008;12;363(1493):947–63.
- [17] Patterson RD, Allerhand MH, Giguère C. Time-domain modeling of peripheral auditory processing: a modular architecture and a software platform. *J Acoust Soc Am* 1995;98(4):1890–4.
- [18] Vuillet E, Magnan A, Ecalle J, Thai-Van H, Collet L. Auditory processing disorder in children with reading disabilities: effect of audiovisual training. *Brain* 2007;130:2915–28.
- [19] Akhoun I, Moulin A, Jeanvoine A, Menard M, Berger-Vachon C, et al. Speech auditory brainstem response (*Speech ABR*) Characteristics depending on recording conditions, and hearing status. An experimental parametric study. *J Neurosci Methods* 2008;175(2):196–205.
- [20] Verpy E, Weil D, Leibovici M, Goodyear RJ, Hamard G et al. Stereocilia-deficient mice reveal the origin of cochlear waveform distortions. *Nature* 2008;13;456(7219):255–8.
- [21] Akhoun I, Tillmann B, Moulin A, Gallégo S, Berger-Vachon C, et al. “2pPP4. Temporal intensity and pitch relationship on speech auditory brainstem response to phonemes/ba/in normal hearing adults.” *POMA* 2008; vol 1 [Epub].
- [22] Bleek S, Ives T, Patterson RD. Aim-mat: the auditory image model in MATLAB. *Acta Acustica United with Acustica* 2004;90:781–8.
- [23] Moore BC, Glasberg BR. A revised model of loudness perception applied to cochlear hearing loss. *Hear Res* 2004;188(1–2):70–88.
- [24] Glasberg BR, Moore BCJ. Derivation of auditory filter shapes from notched-noise data. *Hear Res* 1990;47:103–38.
- [25] Gaumond RP, Molnar CE, Kim DO. Stimulus and recovery dependence of cat cochlear nerve fiber spike discharge probability. *J Neurophysiol* 1982;48(3):856–73.
- [26] Meddis R, Hewitt M. Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification. *J Acoust Soc Am* 1991;89:2866–82.
- [27] Lyon R, Shamma SA. Auditory representation of timbre and pitch. *Springer Handbook of Auditory Research*, (vol. 6) Auditory Computation. In: Haw-

Pour citer cet article : Akhoun I, et al. Approche empirique pour déterminer les générateurs neurophysiologiques sous-jacents des potentiels évoqués auditifs engendrés par des sons de parole (*Speech ABR*). *IRBM* (2010), doi:10.1016/j.irbm.2009.12.005

- kins HL, McMullen T, Popper AN, Fay RR (Eds.), (New-York: Springer) 1996.
- [28] Xu Z, De Vel E, Vinck B, van Cauwenberge PB. Choice of a tone-pip envelope for frequency-specific threshold evaluations by means of the middle-latency response: normally hearing subjects and slope of sensorineural hearing loss. *Auris Nasus Larynx* 1997;24(4):333–40.
- [29] Moushegian G, Rupert AL, Stillman RD. Laboratory note. Scalp-recorded early responses in man to frequencies in the speech range. *Electroenceph clin Neurophysiol* 1973;35(6):665–7.
- [30] Daly D, Roeser R, Moushegian G. The frequency-following response in subjects with profound unilateral hearing loss. *Electroenceph Clin Neurophysiol* 1976;40(2):132–42.
- [31] Faingold CL, Caspary DM. Frequency-following responses in primary auditory and reticular formation structures. *Electroenceph Clin Neurophysiol* 1979;47(1):12–20.
- [32] Hall J. Auditory brainstem frequency following responses to waveform envelope periodicity. *Science* 1979;205(4412):1297–9.
- [33] Greenberg S, Marsh J, Brown W, Smith J. Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones. *Hear Res* 1987;25(2–3):91–114.
- [34] Eggermont JJ, Schmidt PH. The auditory brainstem response. *Evoked Potential Manual*. In: Colon EJ, Visser SL (Eds.), (Amsterdam: Kluwer Academic), 1990; p.41–77.
- [35] Galbraith G, Arbagey P, Branski R, Comerci N, Rector P. Intelligible speech encoded in the human brain stem frequency-following response. *Neuroreport* 1995;6(17):2363–7.
- [36] Krishnan A. Human frequency-following responses: representation of steady-state synthetic vowels. *Hear Res* 2002;166(1–2):192–201.
- [37] Musacchia G, Sami M, Skoe E, Kraus N. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc Natl Acad Sci U S A* 2007;10440 [15894–8 (Epub)].
- [38] Oxerham AJ, Bernstein JG, Penagos H. Correct tonotopic representation is necessary for complex pitch perception. *Proc Natl Acad Sci U S A* 2004;101(5):1421–5.
- [39] Dau T. The importance of cochlear processing for the formation of auditory brainstem and frequency following responses. *J Acoust Soc Am* 2003;113(2):936–50.
- [40] Krishnan A, Swaminathan J, Gandour JT. Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *J Cogn Neurosci* 2008 (in press).
- [41] Rees A, Sarbaz A, Malmierca MS, Le Beau FE. Regularity of firing of neurons in the inferior colliculus. *J Neurophysiol* 1997;77(6):2945–65.
- [42] Voutsas K, Langner G, Adamy J, Odruc M. A brain-like neural network for periodicity analysis. *IEEE Trans Syst Man Cybern B Cybern* 2005;35(1):12–22.
- [43] Kraus N, Bradlow AR, Cheatham MA, Cunningham J, King CD, et al. Consequences of neural asynchrony: a case of auditory neuropathy. *J Assoc Res Otolaryngol* 2000;1(1):33–45.
- [44] Gilbert G, Bergers I, Voüllery D, Lorenzi C. Effects of periodic interruptions on the intelligibility of speech based on temporal fine-structure or envelope cues. *J Acoust Soc Am* 2007;122(3):1336.
- [45] Sheft S, Ardoint M, Lorenzi C. Speech identification based on temporal fine structure cues. *J Acoust Soc Am* 2007;124(1):562–75.
- [46] Lorenzi C, Gilbert G, Carn H, Garnier S, Moore BC. Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc Natl Acad Sci U S A* 2006;103(49):18866–9.
- [47] Gilbert G, Lorenzi C. The ability of listeners to use recovered envelope cues from speech fine structure. *J Acoust Soc Am* 2006;119(4):2438–44.
- [48] Akhoun I, Richard C, Moulin A, Gallégo S, Berger-Vachon C, et al. Comparison of speech vs speech-homologue temporal encoding in human: An auditory brainstem response and categorical perception assessment. *J Acoust Soc Am* 2008;123:3863.