

Génétique et physiologie cellulaire

M^{me} Christine PETIT, membre de l'Institut
(Académie des sciences), professeur

LA VOIX : SA PRODUCTION, SA PERCEPTION

Audition et production vocale sont étroitement liées. Le cours de cette année a porté sur la voix. Depuis deux décennies environ, la recherche portant sur la voix s'est individualisée de celle consacrée à la parole. Physiciens, psycho-physiciens et phoniatres, ont longtemps été les seuls à étudier la voix. Les biologistes commencent à en découvrir l'intérêt ; ils inscrivent tout naturellement leur approche dans une perspective évolutive, et tentent de lier propriétés acoustiques de certaines émissions vocales et propriétés mécaniques des plis vocaux, et ces dernières à l'histologie et à la composition biochimique de cette structure. Le caractère interdisciplinaire de cette recherche se développe, et avec lui se dessine un vrai regain d'intérêt pour ce sujet, dont on apprécie aisément l'importance pour la communication humaine et l'expression artistique.

Le cours a exploré une série de questions : Que sont les sons vocaux, quelles sont leurs caractéristiques acoustiques, du ronronnement aux phonèmes de la parole, au chant, en passant par les cris et les rugissements ? Comment sont-ils produits ? Comment ont-ils évolué dans le règne animal ? Comment a évolué l'appareil vocal qui les produit ? La voix correspond-elle à une catégorie perceptive acoustique, en d'autres termes, fait-elle l'objet d'un traitement spécifique dans le cortex cérébral auditif, comme cela paraît être le cas pour les visages dans le cortex visuel ? Si ce traitement est spécifique, quel est son rôle ? A-t-il un rôle social en contribuant à l'identification des individus ? Quelles informations la voix véhicule-t-elle sur le locuteur, ses caractéristiques physiques, son état psychologique, en particulier émotionnel ? Comment la perception par le locuteur des sons vocaux qu'il émet modifie-t-elle sa propre production de sons ? Quand la boucle audio-phonatoire se met-elle en place chez l'homme ? Que sait-on de son évolution ? Comment les sons environnants, bruits et perception des sons vocaux émis par les autres individus de l'espèce influencent-ils les productions vocales ? Que nous apprend l'étude des anomalies de production des sons vocaux et de perception de la voix observées chez l'homme ?

Nous avons tenté de répondre à ces questions à travers la présentation de données récentes de la littérature scientifique, après avoir introduit les connaissances nécessaires à leur compréhension.

Premier cours. Les productions vocales, du ronronnement à la parole : caractéristiques physiques et physiologie

Le premier cours a porté sur la production de la voix. Il a débuté par un rappel de l'anatomie fonctionnelle du tractus vocal, composé du larynx, organe phonatoire et du tractus supra-laryngé, qui s'étend, chez l'homme, de la glotte aux lèvres, et comporte trois cavités : pharyngée, buccale et nasale. Les éléments du larynx osseux et cartilagineux, les membranes épaisses, ligaments et muscles qui leur sont associés, ainsi que les plis vocaux, et l'espace qu'ils délimitent, la glotte, ont été décrits. Nous avons vu le jeu des muscles qui permet le rapprochement des plis vocaux, condition de leur vibration, ainsi que le rôle de muscles qui tendent ou détendent les plis vocaux, modifiant ainsi leur fréquence de vibration. Les modifications de la configuration du tractus supra-laryngé, à l'origine du langage articulé, ont été présentées.

Le cours a ensuite abordé la production des phonèmes, et plus particulièrement des voyelles. La théorie développée par Gunnar Fant en 1960 (Fant, 1960), connue sous le nom de théorie « source-filtre », postule l'existence de deux éléments distincts impliqués dans la production des voyelles. L'un est une source énergétique : c'est la pression d'un flux d'air provenant des poumons lors de l'expiration ; elle fait vibrer les plis vocaux et crée ainsi des vagues de pression dans l'air expiré, ce qui le sonorise. L'autre élément est un filtre à travers lequel passe le son glottique (ou laryngé). Ce filtre est constitué par le tractus vocal supra-laryngé, dont la géométrie est extrêmement dynamique. Les cavités supra-laryngées changent de forme très rapidement par le mouvement de leurs éléments, par exemple celui de la langue, et leurs résonances se modifient en parallèle. Les voyelles se distinguent des consonnes par leur intensité généralement plus forte, leur durée plus longue, leur stabilité et leur structure spectrale harmonique. Les pics énergétiques du spectrogramme des voyelles sont appelés les « formants ». Certaines fréquences du spectre sonore glottique sont « absorbées » par le tractus vocal supra-laryngé, mais d'autres, dont la fréquence est identique ou voisine de celles des résonateurs supra-laryngés, sont au contraire amplifiées : elles constituent les formants. Ce mode de production de sons vocaux est très général chez les mammifères, même si l'homme en fait un usage spécifique pour former les voyelles. On désigne les formants par la lettre F, de F1 à Fn, en allant des basses aux hautes fréquences, tandis que F0 correspond à la fréquence fondamentale de vibration des plis vocaux. Les formants F1 et F2 sont particulièrement adaptés à la description des voyelles. La fréquence de F1 est comprise entre 300 et 700 Hz, celle de F2 entre 800 et 2 000 Hz. La valeur de F1 est en général corrélée avec l'ouverture de la bouche, et celle de F2, quand elle est élevée, avec une position antérieure de la langue, et quand elle est faible, avec sa position postérieure. L'importance du tractus supra-laryngé dans la reconnaissance de la parole est bien mise en évidence par le chuchotement. Lors du chuchotement, le larynx génère non pas des sons harmoniques, mais un bruit à large bande, qui n'a pas de hauteur tonale (voir plus loin). La parole chuchotée est cependant intelligible en raison de la présence de formants.

En 1972, Kenneth Stevens a proposé une représentation graphique des formants des voyelles, avec sur l'axe vertical les variations de F1, et sur l'axe horizontal celles de F2. Les différentes voyelles forment un triangle dit « vocalique ». Parce que la prononciation des voyelles est caractérisée par la stabilité du conduit vocal pendant leur émission, il est possible de schématiser l'ensemble du tractus supra-laryngé par des parallélépipèdes et des tubes ouverts ou fermés à une extrémité, représentant respectivement ses cavités et ses lieux de constriction. On obtient ainsi une bonne approximation de la fréquence des formants des voyelles correspondantes. La formation des consonnes, hormis les consonnes voisées, repose sur des mouvements complexes et rapides du tractus vocal, qui se caractérisent par l'obstruction du passage de l'air dans les cavités supra-laryngées, et dont la description cinétique reste incomplète. Les consonnes voisées, quant à elles, mettent en jeu la vibration des plis vocaux.

Avant de passer à la physiologie des plis vocaux, leur anatomie et leur histologie ont été décrites. Ils sont composés de fibrocytes qui secrètent une matrice extracellulaire, dont la composition varie avec le stress mécanique. Ce tissu formé de fibres élastiques et de glycosaminoglycanes, dont le composé le plus abondant est l'acide hyaluronique, confère des propriétés viscoélastiques particulières aux plis vocaux. Quand la viscosité de ce tissu augmente, sa mise en vibration comme le maintien de cette vibration sont plus difficiles ; quand la rigidité des plis vocaux diminue, la stabilité fréquentielle de leurs vibrations aussi. De façon surprenante, on est loin d'avoir une compréhension intégrée de la composition biochimique, de la structure histologique, des propriétés biophysiques et des caractéristiques physiologiques des plis vocaux. On peut soupçonner la présence de plusieurs types de mécanorécepteurs dans les plis vocaux, mais aucun n'a encore été identifié. Les plis vocaux sont soumis à des tensions par des muscles qui agissent selon un axe antéro-postérieur. Ils sont soumis à des vibrations induites par la pression sous-glottique ; l'amplitude de leurs vibrations est corrélée à cette pression.

Quelques éléments essentiels de la physiologie des plis vocaux ont ensuite été présentés : leur mode et leur fréquence de vibration. Le déplacement des plis vocaux peut se faire selon plusieurs régimes. Durant la phonation des voyelles, les plis vocaux sont le siège d'une oscillation auto-entretenu. La théorie la plus communément admise pour décrire cette oscillation est la théorie dite myo-élastique aérodynamique. Elle stipule que la soufflerie pulmonaire provoque une vibration des plis vocaux quand elle parvient à vaincre la résistance que l'élasticité des plis vocaux oppose au passage de l'air. Les interprétations de cette oscillation auto-entretenu font appel à la loi de conservation de l'énergie adaptée aux fluides par le théorème de Bernoulli, qui décrit la variation, en sens inverse, de la vitesse de déplacement des particules de fluides et de la pression locale. Une approximation du mouvement des plis vocaux par celui d'un simple oscillateur harmonique a été initialement proposée. L'analyse laryngoscopique du mouvement des plis vocaux, aujourd'hui réalisée à l'aide d'une caméra vidéo ultra rapide, démontre qu'il est non-uniforme et beaucoup plus complexe qu'on le pensait. D'une part, la partie superficielle des plis vocaux (enveloppe) peut se déplacer de façon quasi indépendante de sa partie profonde (corps). D'autre part, l'enveloppe des plis vocaux est elle-même le siège de mouvements différents sur sa face trachéale (basale) et sur sa face pharyngée (apicale).

Le mouvement des plis vocaux se décompose en plusieurs étapes. Dans la première, sous l'effet de la contraction musculaire, il y a adduction des plis vocaux, la glotte est quasiment fermée, et les plis vocaux sont sous tension. La pression de l'air sous-

glottique augmente. Lorsqu'elle atteint un certain seuil, elle va vaincre les forces élastiques des plis vocaux, qui s'ouvrent sur leur face basale et adoptent ainsi une configuration dite convergente. Dans cette configuration, la pression intra-glottique augmente. Les plis vocaux commencent à s'écarter dans leur partie apicale. L'écartement se poursuit tant que les forces aérodynamiques venues de la trachée sont supérieures aux forces élastiques qu'opposent les plis vocaux. L'augmentation du débit intra-glottique qui s'ensuit s'accompagne d'une baisse de la pression intra-glottique (théorème de Bernoulli). La glotte, qui était ouverte de façon équivalente dans la partie apicale et basale des plis vocaux, passe alors en configuration divergente par un mouvement de fermeture de la partie basale des plis vocaux. La pression intra-glottique devient trop faible pour s'opposer aux forces élastiques des plis vocaux : la glotte se referme. Il s'ensuit une augmentation de la pression sous-glottique et le cycle reprend... (d'après I.R. Titze). Parce que la vibration des plis vocaux est non-linéaire, à sa fréquence fondamentale est associé un large spectre d'harmoniques. Certaines de ces fréquences seront amplifiées par résonance lors du passage de l'air à travers le tractus supra-laryngé. La fréquence fondamentale de vibration des plis vocaux donne la perception de la hauteur tonale de la voix : elle diminue de l'enfant, chez lequel elle se situe entre 250 et 400 Hz (500 Hz pour le cri) à l'adulte, chez lequel elle est de l'ordre de 200 Hz pour les femmes et 125 Hz pour les hommes. En voix chantée, la fréquence fondamentale peut atteindre 1,5 kHz chez les femmes. La fréquence fondamentale de vibration des plis vocaux permet d'identifier la hauteur tonale de la voix, qui contribue à la reconnaissance du locuteur. Elle transmet également des informations sur l'état émotionnel du locuteur. La détection de la hauteur tonale de la voix est aussi essentielle pour séparer ou au contraire rattacher à une même source sonore des messages acoustiques. Dans les langues tonales, elle porte en outre une information linguistique.

La dernière partie de ce cours a porté sur la descente du larynx. Elle conditionne la création de la cavité pharyngée et la libération de la langue, qui devient alors très mobile, propriété essentielle à l'élaboration d'une parole articulée. De fait, chez l'homme, la mise en place du langage durant l'enfance s'accompagne de la descente du larynx. C'est dans le cadre des interrogations qui portent sur l'origine du langage que les changements évolutifs de l'anatomie du tractus vocal ont été initialement étudiés. La paléoanthropologie a longtemps été la seule voie d'approche. Elle se heurte cependant à une difficulté majeure ; le larynx ne comporte qu'une seule structure qui peut se fossiliser : l'os hyoïde. Parce que l'on admet que la verticalisation progressive de la base du crâne chez les hominidés a joué un rôle essentiel dans la descente du larynx, elle est considérée comme un indice de cette dernière. Aujourd'hui, l'étude des génomes des hominidés offre la possibilité, lorsque l'on identifie un gène impliqué dans la production du langage articulé chez l'homme, de s'interroger sur son émergence durant l'évolution (voir quatrième cours). Depuis environ deux décennies, une autre approche s'est développée, dont l'objectif est de tracer un schéma évolutif plausible de l'appareil phonatoire. Elle conduit à mettre en lumière un ensemble d'adaptations de l'appareil laryngé et supra-laryngé, dont on essaie de rendre compte par les avantages sélectifs qu'elles confèrent. Nous avons illustré cette nouvelle démarche par trois exemples.

Le mode de production de sons vocaux de basse fréquence chez l'éléphant est inconnu. Correspond-il à un mode de vibration des plis vocaux qui répond à la théorie myo-élastique dynamique (dans laquelle la fréquence fondamentale de vibration des plis vocaux est proportionnelle à leur longueur [en unité logarithmique]),

ou bien est-il semblable à celui qui sous-tend le ronronnement des félins ? Ce ronronnement est dû à un son laryngé qui a pour origine une modulation régulière du flux aérien venant des poumons par une activité intermittente des muscles laryngés. Ces muscles se contractent avec une fréquence de 20 à 30 Hz. Elle peut atteindre 200 Hz chez certains félins; elle est de l'ordre de 100 Hz chez le tigre. Pour répondre à cette question, les expérimentateurs ont utilisé un larynx d'éléphant excisé (Herbst *et al.*, 2012). Dans cette préparation, la commande neuronale des muscles laryngés est perdue. Les expériences consistent donc à explorer la possibilité d'obtenir la production de sons de basse fréquence par le mécanisme aérodynamique myo-élastique sollicité par l'envoi d'air sous pression au-dessous des plis vocaux. Les sons produits sont enregistrés par un microphone, les mouvements des plis vocaux sont suivis par une caméra ultrarapide et les variations de la surface de contact des plis vocaux sont appréciées par électroglottographie. Ce dispositif a permis de montrer que, sous l'effet de la pression de l'air, on peut obtenir une oscillation auto-entretenu des plis vocaux, qui conduit à l'émission de sons de basse fréquence. Ces résultats n'excluent pas l'existence d'un mécanisme du type ronronnement chez l'éléphant, mais montrent qu'il n'est pas nécessaire pour rendre compte de la production de ses vocalisations de basse fréquence.

Les travaux menés chez l'animal ont révélé que la position basse du larynx n'est en rien une particularité des hominidés. On la retrouve chez d'autres mammifères comme le koala, ou le cerf rouge. Chez ce dernier, le brame est composé de sons plus graves que ceux émis habituellement. On observe alors que le larynx, dont on peut suivre le déplacement par celui de son cartilage thyroïde, saillant sous la peau, descend d'une quarantaine de centimètres (Fitch & Reby, 2001). Il s'ensuit un accroissement considérable de la cavité pharyngée, qui explique l'enrichissement du spectrogramme des émissions vocales en formants de basse fréquence. Quel avantage sélectif pourrait être associé à l'accroissement de la cavité pharyngée ? Au sein d'une même espèce, la fréquence des formants, parce qu'elle est proportionnelle au volume des cavités supra-laryngées, est un indicateur du gabarit de l'animal, d'où l'hypothèse dite « hypothèse de l'exagération ». La descente fonctionnelle du larynx au cours du brame conférerait aux mâles une présentation avantageuse de leurs qualités physiques, qui, de plus, échappe à une possible pondération par la vision, puisque le brame a essentiellement lieu la nuit. Cette présentation avantageuse attirerait les femelles, tout en éloignant les mâles compétiteurs. Une descente fonctionnelle du larynx existe aussi chez le lion, le tigre, ou la panthère, mais son mécanisme est différent. Au total, ces travaux indiquent que la descente du larynx n'est pas un événement évolutif unique qui aurait permis l'apparition du langage articulé. Elle serait survenue plusieurs fois durant l'évolution en raison de l'avantage sélectif qu'elle confère lors la reproduction, et par convergence comme le suggère la diversité des mécanismes impliqués. Cette descente du larynx est considérée par certains comme une préadaptation au langage. Chez les chimpanzés comme chez l'homme, on observe chez les tout-petits une baisse de la position du larynx (Nishimura *et al.*, 2003). Cette première étape de descente du larynx est parfois considérée comme un moyen de sécuriser le processus de déglutition en rendant le larynx plus indépendant des mouvements de la langue. Il faut rappeler que le rôle du larynx est à l'origine la protection des voies aériennes contre le passage de liquides ou de particules, alimentaires ou autres. Cette étape serait donc une préadaptation à la mise en place de la phonation chez les primates. Elle est complétée, chez l'homme, par une autre étape de descente du larynx, qui n'existe

pas chez le chimpanzé, et qui serait le résultat du changement de position de la base du crâne. Chez l'homme, une descente additionnelle du larynx survient chez le garçon à la puberté ; elle n'accroît pas les possibilités phonétiques. Il s'agit d'un caractère sexuel secondaire, par lequel la voix devient plus grave.

Pour terminer, deux exemples tirés de la littérature scientifique récente, qui illustrent la plasticité de la production vocale chez les mammifères, ont été exposés. Une éléphant d'Asie, élevée dans un zoo en Corée du Sud, a appris à prononcer quelques mots de coréen. Pour produire les formants de certaines voyelles, elle introduit sa trompe dans sa cavité buccale, ce qui modifie sa fréquence de résonance et lui permet d'émettre des formants de fréquence appropriée (Stoeger *et al.*, 2012). Deuxième exemple. Dans le chant soprano, lorsque la fréquence fondamentale de vibration des plis vocaux s'élève, elle devient trop élevée pour tirer partie d'un ensemble de résonateurs du tractus vocal dont la fréquence de résonance est plus basse ; en conséquence l'intensité de la voix ne peut être forte. Un ajustement de la fréquence de résonance de la cavité buccale et de la fréquence fondamentale de vibration des plis vocaux est possible, par l'élaboration d'un sourire et par l'abaissement de la mâchoire. Des chercheurs japonais (Koda *et al.*, 2012) se sont demandés si le chant soprano était une spécificité humaine. Ils se sont intéressés au chant du gibbon. La femelle peut produire un son dont la fréquence fondamentale est très intense : en conséquence, soit la fréquence de résonance du tractus vocal chez le gibbon n'est pas indépendante de la fréquence fondamentale de vibration des plis vocaux, soit le gibbon chante comme une soprano. En modifiant artificiellement la fréquence de résonance des cavités du tractus supra-laryngé (en faisant respirer à l'animal de l'air mélangé à de l'hélium), ils ont montré que les gibbons utilisent la même technique vocale que les sopranos pour produire des sons vocaux de fréquence élevée et de forte intensité.

Deuxième cours. La perception de la voix : son traitement comme objet acoustique, aspects ontogéniques et évolutifs

Le deuxième cours a porté sur la perception de la voix, du traitement de ses caractéristiques physiques jusqu'à sa représentation comme identifiant de l'individu. L'idée selon laquelle à la voix pourrait être associée une catégorie perceptive auditive spécifique s'ancre sur les travaux effectués sur la perception du visage. En d'autres termes, la perception de la voix et celle du visage, éléments importants de la communication sociale, bénéficieraient d'un traitement spécifique. La perception des visages, comme celle de la voix, contribue à l'identification de l'individu, de son genre, son âge, son état de santé, ses émotions et ses intentions. Les travaux réalisés sur la perception des visages depuis les années 1970 ont très largement inspiré ceux menés sur la reconnaissance de la voix vingt ans plus tard. Quelques résultats essentiels portant sur la reconnaissance des visages ont été exposés. Trois questions ont été évoquées. Comment une vision holistique des visages est-elle générée ? Quelles sont les aires cérébrales impliquées ? Peut-on parler d'aires cérébrales spécifiquement dédiées à la reconnaissance des visages ?

La reconnaissance holistique des visages a été démontrée par de nombreux tests. Sa sensibilité au sens haut/bas de la présentation du visage (qui n'intervient pas dans la reconnaissance des objets) (Little *et al.*, 2011b ; Little *et al.*, 2011a), aux rapports entre les divers attributs du visage, les yeux, le nez, la bouche (Rossion, 2008) et

enfin, aux contrastes de luminosité existant entre les diverses parties du visage, ont été décrits. L'imagerie fonctionnelle cérébrale par résonance magnétique nucléaire (IRMf) a montré l'implication d'une aire cérébrale activée lors de la présentation d'un visage : l'aire faciale fusiforme (abrégée en langue anglaise par FFA). Sont également impliquées deux autres aires : l'aire faciale occipitale et une toute petite aire du lobe temporal inférieur, dans sa partie antérieure. La réponse de ces trois régions à la présentation d'un visage immobile est plus importante dans le cortex droit que dans le cortex gauche. La région cérébrale qui est activée par l'expression des visages se situe, quant à elle, dans le sillon temporal supérieur (STS). Les arguments en faveur d'un traitement cortical spécifique des visages proviennent de l'étude d'une situation pathologique, la prosopagnosie. Elle se manifeste par une absence de confiance dans la reconnaissance des visages familiaux. Les prosopagnosies traumatiques indiquent, elles aussi, la dominance de l'hémisphère droit dans la reconnaissance des visages. Elles montrent qu'en cas de lésion, il n'y a pas de compensation possible par une autre aire cérébrale. La prosopagnosie peut être héréditaire. Elle est alors transmise sur le mode autosomique dominant ; aucun gène responsable n'a cependant encore été identifié. Certaines données de la littérature scientifique font état d'une fraction très importante des individus qui seraient atteints de prosopagnosie : le chiffre de 2,3 % de la population est avancé, soit un chiffre voisin de celui rapporté pour la dyslexie ou la dyscalculie. Si les bases moléculaires de la prosopagnosie ne sont pas connues, des éléments de son substrat anatomique ont cependant été identifiés dans une étude récente (Thomas *et al.*, 2009). En utilisant la tractographie, l'altération de deux faisceaux du cortex occipito-temporal ventral a été observée. L'un est le faisceau longitudinal inférieur et l'autre, le faisceau fronto-occipital inférieur. Ces deux faisceaux traversent l'aire faciale fusiforme. Ils sont affectés dans les hémisphères droit et gauche. Cette étude ne porte que sur six individus atteints de prosopagnosie, et ses conclusions demandent à être confirmées. Certains continuent de douter de l'existence de zones cérébrales dédiées au seul traitement holistique du visage. Ils avancent l'idée d'une très haute expertise visuelle spatiale associée à l'aire faciale fusiforme qui serait, depuis la naissance, entraînée à la reconnaissance des visages. À l'appui de cette proposition, un article récent rapporte l'activation de cette région cérébrale chez des individus experts dans la reconnaissance des voitures ou des avions, auxquels des photos de ces véhicules sont présentées (McGugin *et al.*, 2012).

La recherche des centres corticaux de reconnaissance de la voix repose sur l'imagerie fonctionnelle cérébrale. L'IRMf se heurte à une difficulté d'utilisation en raison du bruit des machines : il est de l'ordre de 60 dB, et impose donc des stimulations sonores d'intensité supérieure, en règle générale de 80 à 90 dB. On a aussi recours à l'imagerie fonctionnelle en infrarouge proche, mais elle ne détecte que l'activité des couches superficielles du cortex. La recherche de ces centres corticaux a débuté en 2000 par une étude en IRMf menée chez l'adulte (Belin *et al.*, 2000). Les réponses corticales à des sons vocaux et non vocaux ont été comparées. L'activité liée aux sons vocaux était répartie en trois points du STS : points antérieur, central et postérieur (le point central étant proche du gyrus de Heschl, et la région postérieure s'étendant dorsalement en direction du *planum temporale*). La réponse de l'hémisphère droit était plus importante que celle de l'hémisphère gauche. D'autres expériences ont utilisé des « sons témoins » plus adaptés : des sons harmoniques produits par une cloche ou des bruits blancs, modulés en amplitude comme les sons vocaux de l'expérience. Elles ont confirmé

l'activation du STS droit par la voix, et montré que la réponse venait principalement de sa région centrale. Cette aire corticale a été nommée aire vocale temporale (abrégée en anglais par VTA). Elle est surtout sensible au contenu fréquentiel de la voix, et peu à ses variations en amplitude. Une approche complémentaire a été utilisée, qui a conforté ces résultats. Elle a consisté en l'application de filtres semblables sur des sons vocaux et des sons non vocaux harmoniques. Un parallélisme a été observé entre l'intensité de l'activation de la région VTA et l'aptitude à la reconnaissance de la voix. La baisse de l'activité des régions VTA, suite au filtrage de certaines fréquences, s'accompagne d'une baisse de la reconnaissance de la voix du locuteur. La reconnaissance du locuteur par sa voix repose sur la détection de sa fréquence fondamentale F0. La fréquence fondamentale peut cependant être absente du spectre de parole pour diverses raisons : absence de résonance des cavités du tractus vocal à cette fréquence ou atténuation voire non transmission des basses fréquences (par le téléphone par exemple). Le système auditif central peut cependant l'extraire à partir de ses harmoniques : on parle alors de hauteur tonale virtuelle. L'extraction de la fréquence fondamentale manquante existe chez les poissons, les oiseaux et les mammifères. Cependant l'absence de fréquence fondamentale dans le spectre modifie le timbre.

Au-delà de ces corrélations, l'établissement d'un lien de causalité entre l'activation de la région VTA et la reconnaissance de la voix a été recherché. On ne connaît pas ce qui serait l'équivalent de la prosopagnosie, soit un déficit isolé de la reconnaissance de la voix. On a pensé que les amusies, qui sont caractérisées par une incapacité à reconnaître un changement de tonalité dans une suite mélodique, pouvaient s'accompagner de troubles de la reconnaissance de la voix, mais il n'en est rien (Tillmann *et al.*, 2012). Pour tester ce lien de causalité entre l'activité de la région VTA de l'hémisphère droit et la reconnaissance des voix, une stimulation trans-crânienne, ciblée sur cette région corticale pour en inhiber l'activité, a été effectuée chez plusieurs individus (Bestelmeyer *et al.*, 2011). Ces expériences démontrent que lorsque la stimulation trans-crânienne est ciblée sur les régions antérieure et centrale du STS droit, la reconnaissance des voix décroît de façon significative. Cette dernière n'est pas affectée par une stimulation trans-crânienne portant sur une autre région du cerveau. De plus, ciblée sur la région VTA, elle n'affecte pas la discrimination d'intensité sonore. D'autres expériences apportent des arguments en faveur des conclusions précédentes. Ainsi par exemple, on peut suivre l'adaptation d'une réponse corticale par la baisse de l'activité de la région correspondante. Dans un test où des syllabes différentes sont prononcées par une même voix, il y a adaptation de la réponse de la région VTA de l'hémisphère droit, alors que si la même syllabe est prononcée par 12 locuteurs différents, il n'y a pas d'adaptation localisée dans cette région (Belin & Zatorre, 2003). Plus récemment, des travaux d'imagerie cérébrale fonctionnelle (Formisano *et al.*, 2008) ont permis, par le développement d'algorithme d'analyse des réponses corticales, de définir plus finement les régions cérébrales de la reconnaissance de la voix en corrélation avec sa fréquence fondamentale, et celles de la reconnaissance des voyelles en corrélation avec la valeur des formants F1 et F2 qui les caractérisent. Quel rapport a l'aire VTA avec les cartes tonotopiques cérébrales ? La mise en évidence des cartes tonotopiques corticales par IRMf chez l'homme s'est heurtée à l'utilisation de stimulations sonores de forte intensité, qui étend la région cochléaire stimulée hors de la stricte région dévolue à la détection de la fréquence du son incident. La sensibilité de la technique a été récemment accrue par l'utilisation d'une stimulation sonore dont la

fréquence augmente de façon continue (Striem-Amit *et al.*, 2011). Six cartes tonotopiques (dont l'organisation est comme celle des cartes rétinotopiques, en miroir) ont pu être détectées, dont 4 dans le cortex temporal. L'emplacement de ces cartes suggère leur présence au niveau du VTA.

Parce que la perception de la hauteur tonale de la voix s'effectue souvent à partir de ses harmoniques, un travail récent (Oxenham *et al.*, 2011), qui bouscule la vision antérieure de leur perception, a été brièvement exposé. On tient généralement pour acquis le fait que la reconnaissance de la hauteur sonore repose sur la décharge cadencée (en anglais, *phase locking*) des neurones auditifs par rapport à l'onde sonore. De fait, chez l'homme, la reconnaissance des mélodies est sévèrement dégradée au-delà de 4 à 5 kHz, valeurs au-delà desquelles ce codage temporel de la fréquence du son par celle des décharges neuronales n'existe plus. Les fréquences les plus hautes des instruments d'orchestre, telles que celles du piano ou du piccolo, ne se situent pas au-delà de 5 kHz, d'où l'idée que c'est la contrainte du codage temporel périphérique de la fréquence sonore qui explique la limite de notre perception musicale. Ce travail démontre qu'un accord harmonique dans lequel aucune des fréquences n'est inférieure à 5 kHz permet une excellente détection de la hauteur tonale. Même lorsque la première harmonique du spectre atteint une valeur de 12 kHz, il y a encore une détection satisfaisante de la hauteur tonale. Ces résultats indiquent que le système auditif central peut extraire la fréquence fondamentale manquante à partir d'une information périphérique de nature strictement tonotopique. Pourquoi la possibilité de déceler une hauteur tonale à partir d'accords harmoniques de fréquences si élevées a-t-elle jusque-là échappé aux différents investigateurs ? L'explication se tient vraisemblablement dans les caractéristiques des sons utilisés. Habituellement, ils ne comportent que 3 ou 4 harmoniques ; dans ce travail, ils en comportent une douzaine, auxquelles est associé un bruit masquant d'environ 50 dB qui pouvait supprimer la perception des produits acoustiques de distorsion et rendre ainsi les sons davantage perceptibles.

Enfin, le rôle possible de la reconnaissance des phonèmes dans celle de la voix a été approché à travers une première étude (Perrachione *et al.*, 2011), qui montre que l'intégration du contenu linguistique accroît la reconnaissance du locuteur par sa voix. La complexité du monde social allant grandissant, elle exercerait une pression de sélection sur les mécanismes cérébraux pour qu'ils intègrent les différentes sources d'information provenant de l'environnement.

Troisième cours. Le contrôle de la production vocale par la perception auditive : la « boucle » audio-phonatoire

Le troisième cours a porté sur les trois points suivants :

- l'intégration de la reconnaissance des voix et des visages, qui contribue à accroître celle des personnes ;
- l'ontogénie de la reconnaissance des paramètres acoustiques vocaux chez l'homme.
- le contrôle de la production vocale *via* la perception auditive, ou boucle audio-phonatoire.

L'intégration sensorielle bimodale voix-visage a été approchée par l'imagerie fonctionnelle cérébrale (von Kriegstein & Giraud, 2006). Ces travaux ont eu recours à une approche dite d'« amorçage implicite ou inconscient » (en anglais, *implicit*

priming). Elle consiste en la présentation répétée de deux types de signaux sensoriels associés. On évalue ensuite son effet sur la réponse à la stimulation à un seul de ces deux types de signaux. Ainsi après une période d'amorçage implicite qui porte sur l'association voix-visage de locuteurs, une stimulation exclusive par la voix fait apparaître une activation des aires cérébrales de réponse au visage. Des tests comportementaux conçus selon le même protocole montrent que la reconnaissance de la voix est améliorée par cet amorçage. Cet effet est bien moindre si l'amorçage porte sur l'association voix-visualisation du nom écrit du locuteur. Récemment, un travail (Blank *et al.*, 2011) s'est attaché à rechercher les connexions entre les aires cérébrales impliquées dans la reconnaissance du visage et celles impliquées dans la reconnaissance de la voix, par tractographie. Après définition par IRMf de l'aire FFA de reconnaissance des visages chez différents individus et sa division en deux sous-régions, l'une, vFFA, strictement visuelle, et l'autre, cFFA (pour *crossmodality* FFA), qui peut être activée par la voix seule après un amorçage portant sur l'association voix-visage, leurs connexions avec les aires de reconnaissance de la voix dans le STS ont été recherchées. Trois faisceaux qui associent l'aire FFA aux trois sous-régions du STS (antérieure, centrale, postérieure) ont été mis en évidence. Les faisceaux les plus développés s'étendent entre l'aire cFFA et l'aire de reconnaissance de la voix dans le STS, dans sa partie moyenne et antérieure. L'existence de ces faisceaux suggère une interaction directe entre les aires de reconnaissance de la voix et celles dévolues à la reconnaissance des visages, qui optimise la reconnaissance de l'identité du locuteur. La reconnaissance parole-visage du locuteur qui s'exprime (présentation audio-vidéo) a aussi été étudiée. Dans un article publié en 2008 (von Kriegstein *et al.*, 2008), les auteurs rapportent que la perception de la mobilité des visages implique une aire située dans le STS. Cette aire n'est pas activée par une présentation voix-visage ; elle l'est en revanche par la parole seule après une présentation préalable de l'association parole-visage du locuteur qui s'exprime (montage audio-vidéo), tandis que la région FFA ne l'est pas. Les mêmes investigations menées chez des sujets prosopagnosiques ont montré que chez eux, la reconnaissance de la parole tire parti de la perception de la mobilité du visage, mais que la reconnaissance du locuteur par sa voix après une présentation de l'association voix-visage n'est pas améliorée. Ainsi, deux réseaux neuronaux différents pourraient être impliqués dans la communication sociale qui se fonde sur l'intégration des informations audiovisuelles susmentionnées. L'un tirerait parti de l'information portée par la configuration stable du visage du locuteur pour reconnaître sa voix, l'autre, de la mobilité de son visage pour améliorer la perception de sa parole. L'activité du premier passe au niveau visuel par celle de la région FFA, avec une dominance de l'hémisphère droit. L'activité du second implique au niveau visuel celle des aires du STS de l'hémisphère gauche.

La deuxième partie du troisième cours a porté sur l'ontogénie de la perception des différents paramètres acoustiques de la voix. Pour explorer la perception acoustique chez le nouveau-né et le jeune enfant, on peut avoir recours à « la négativité de discordance » (Pena *et al.*, 2012). Il s'agit d'un potentiel électrique déclenché par un événement. Il a été découvert en 1978 par le Finlandais Risto Näätänen (Naatanen *et al.*, 1978). Il traduit pour l'essentiel une réponse automatique et inconsciente à tout type de discordance introduite dans un stimulus sensoriel régulier. Le sujet est mis en présence d'un signal répété dont la monotonie est rompue par la modification du paramètre dont on veut apprécier s'il est ou non décelé (voir aussi cours de l'année 2011/2012). Ce potentiel est enregistré par des

électrodes placées sur le scalp. On peut également détecter la négativité de discordance par magnétoencéphalographie. La magnétoencéphalographie mesure les champs magnétiques induits par l'activité électrique des neurones du cerveau. Elle permet notamment d'étudier la réponse du fœtus *in utero*.

Nous avons vu qu'*in utero*, le fœtus est sensible aux traits prosodiques de sa langue maternelle. Dans une étude publiée en 2009 (Mampe *et al.*, 2009), les cris de nouveau-nés de parents parlant le français et de nouveau-nés de parents parlant l'allemand ont été comparés. Leurs spectrogrammes vocaux différaient : le temps mis par le cri pour atteindre sa fréquence et son intensité maximales était plus long pour les nouveau-nés de parents s'exprimant en français, et la baisse en intensité et en fréquence plus lente pour les nouveau-nés de parents parlant l'allemand. Or, ces caractéristiques correspondent précisément à la prosodie de leurs langues maternelles respectives. Donc, non seulement les nouveau-nés ont mémorisé les profils d'intonation de leurs langues maternelles, mais ils sont capables de les reproduire. Le cri du nouveau-né ne reproduit que les contours mélodiques de la parole. Or ces derniers font principalement appel aux mécanismes respiratoires et laryngés et pas à la maturité articulatoire. L'IRMf a montré que chez le nouveau-né, en réponse à la musique, c'est déjà l'hémisphère droit qui est préférentiellement activé, alors que les deux hémisphères sont activés par les sons de parole. Des études tirant parti de la négativité de discordance ont permis de montrer que la détection de la fréquence fondamentale se met en place entre le troisième et le quatrième mois après la naissance (He & Trainor, 2009), celle de l'harmonicité autour du sixième mois (Folland *et al.*, 2012).

Puis nous avons examiné le rôle de l'environnement sur la reconnaissance des phonèmes et la qualité des vocalisations produites chez le tout petit. Avant six mois, l'enfant reconnaît les catégories phonétiques de sa langue maternelle. Entre six et huit mois, les enfants sont capables de segmenter les mots dans la parole en détectant la probabilité de transition entre les syllabes. Ils extraient la régularité arithmétique de la combinaison de syllabes dans les phrases. À neuf mois, l'enfant est sensible aux règles phono-tactiques des mots, qui sous-tendent la probabilité de survenue d'une séquence phonétique. Avant la fin de la première année, l'exposition de l'enfant à sa langue maternelle a modifié la façon dont il perçoit une autre langue. L'exposition à la langue maternelle réduit l'aptitude de l'enfant à discriminer les sons de parole d'une langue étrangère. Peut-on modifier cette perte de sensibilité aux sonorités d'une langue étrangère chez l'enfant ? Des enfants dont la langue maternelle est l'anglais ont été ou non exposés directement à un locuteur s'exprimant en chinois mandarin pendant une période de quatre semaines (12 sessions d'exposition de 25 minutes) (Kuhl *et al.*, 2003). Les enfants ont été entraînés à tourner la tête dès qu'ils perçoivent un changement dans une séquence sonore. Les résultats montrent une bien meilleure perception des sons de la langue étrangère chez les enfants qui ont été exposés au locuteur s'exprimant en chinois mandarin. Les mêmes tests, dans lesquels les enfants sont exposés au chinois mandarin, soit prononcé par une personne présente, soit provenant d'un enregistrement audio, montrent que l'allongement de la période de sensibilité à la langue étrangère requiert le contact avec une personne qui parle cette langue.

Les interactions sociales jouent aussi un rôle dans l'étendue du répertoire vocal. Un travail effectué chez les enfants de sept mois (Goldstein *et al.*, 2003) montre que la présence de la mère, dès lors qu'elle manifeste un comportement social positif même sans expression vocale, modifie la qualité de leur babillage : son contenu en syllabes croît. Les interactions sociales pèsent donc sur la qualité à la fois de la perception et de l'expression vocales.

Des travaux mesurant la réponse des aires cérébrales chez l'enfant par spectroscopie en infrarouge proche ont montré l'activation d'aires cérébrales sensibles à la voix chez l'enfant de sept mois (Grossmann *et al.*, 2010). La comparaison des réponses à une prosodie neutre, joyeuse, ou qui marque la colère, a permis de mettre en évidence une réponse différentielle des différentes aires sensibles à la voix selon le contenu émotionnel de la parole. Les résultats indiquent que dès cet âge, une prosodie comportant un contenu émotionnel est traitée dans l'hémisphère droit, dominance que l'on retrouve chez l'adulte (voir plus haut). La colère induit une activité cérébrale plus importante que l'expression joyeuse. Des études comparables ont ensuite été menées par IRMf. Elles ont révélé l'activité de régions cérébrales sensibles à la voix dès le troisième mois de vie (région STS de l'hémisphère droit) et la perception du contenu émotionnel de la voix dès cet âge, mais limitée aux expressions tristes (activation du cortex orbito-frontal et de l'insula) (Blasi *et al.*, 2011). Qu'il s'agisse de colère ou d'expression triste, l'interprétation donnée suppose que la priorité dans le développement est donnée à leur perception, situation interprétée comme la priorité donnée à une information négative, car plus essentielle à la survie qu'une information positive. La tractographie utilisée pour explorer la mise en place de la connectivité inter- et intra-hémisphérique a révélé (Perani *et al.*, 2011) de nettes différences entre le nouveau-né et l'adulte ; chez le nouveau-né, les connections inter-hémisphériques sont très bien développées, mais les faisceaux intra-hémisphériques le sont très peu. Le faisceau qui connecte le cortex temporal à l'aire de Broca (gyrus frontal inférieur), *via* le faisceau arqué et le faisceau longitudinal supérieur, est absent à la naissance. En revanche, l'autre faisceau dorsal qui connecte le cortex temporal supérieur, dans sa partie moyenne et postérieure, au cortex pré-moteur, et qui passe par les deux mêmes faisceaux que le précédent, est présent dès la naissance. Ce faisceau, myélinisé plus précocement, correspondrait à une boucle sensori-motrice qui permettrait le contrôle de l'expression vocale dès les premiers mois de vie. C'est cette boucle de contrôle audio-phonatoire qui a été étudiée dans la troisième partie de ce cours.

En 1910, Étienne Lombard, otorhinolaryngologiste français, dans une communication orale à l'Académie de médecine, rapporte l'observation qu'il a faite chez un patient à l'hôpital Lariboisière. Sa communication est intitulée « le signe de l'élévation de la voix ». L'effet qu'il y rapporte, l'augmentation de l'intensité de la voix déclenchée par celle du bruit ambiant, sera ensuite nommé « effet Lombard » quand la découverte commencera à être connue, et surtout quand aura cessé la querelle qui oppose Lombard au hongrois Bárány, qui revendique l'antériorité de la découverte. Celle-ci sera finalement attribuée à Lombard. Lombard conclut au rôle « de l'oreille acoustique dans la régulation de l'intensité de la voix ». Ses études ultérieures montreront qu'il ne s'agit pas d'un réflexe. Ce rétrocontrôle de la production vocale par l'audition est une boucle sensori-motrice. L'intérêt pour l'effet Lombard va croissant depuis une vingtaine d'années, comme le montre le nombre de publications annuelles sur ce thème (Brumm & Zollinger, 2011). Il est présent chez les mammifères (singes-écureuil, chat, baleine et chauves-souris) et les oiseaux, et est décelé lors de la production de vocalisations aussi bien innées qu'acquises. Son existence chez les batraciens est débattue. Le regain d'intérêt qu'il suscite tient aux modifications de l'environnement : les concentrations humaines croissantes dans les villes, et pour les oiseaux, le déboisement qui les expose aux bruits urbains intenses. Il tient au fait que l'effet ne se limite pas à une simple élévation de l'intensité de la voix dans les milieux bruyants ; la boucle audio-

phonatoire s'étend à d'autres paramètres acoustiques de la voix, notamment sa fréquence fondamentale et celle des formants. Des expériences fondées sur la négativité de discordance ont montré que dans cette boucle audio-phonatoire, des sons harmoniques vocaux produits par l'auditeur étaient bien plus efficaces que ceux provenant d'une écoute passive, indiquant que le traitement sensoriel dans la boucle audio-phonatoire est amélioré par des mécanismes moteurs. Les sons riches en harmoniques sont bien plus efficaces que les sons purs (Behroozmand *et al.*, 2011).

Une incursion a alors été faite dans le domaine de la réponse aux sons harmoniques par la présentation de la découverte de neurones qui répondent aux harmoniques, réalisée en 2005 (Bendor & Wang, 2005). Ces neurones, enregistrés chez le marmouset, répondent à la fois aux sons purs et à des sons harmoniques dépourvus de leur fréquence fondamentale, mais dont la fréquence fondamentale est la même que celle des sons purs. La réponse de ces neurones est très robuste : elle persiste même lorsqu'un grand nombre d'harmoniques sont supprimées. Ces neurones se situent en bordure du cortex auditif central et au niveau de la ceinture latérale antérieure et médiane ; la fréquence fondamentale à laquelle ils répondent n'excède pas 800 Hz (en cohérence avec les fréquences des vocalisations dans cette espèce). L'ontogénie de la boucle audio-phonatoire permet d'apprécier le rôle qu'elle pourrait jouer dans l'apprentissage du langage, en permettant à l'individu de procéder par essai-erreur-correction. Dans un travail récent (MacDonald *et al.*, 2012), l'hypothèse du rétrocontrôle de la production des voyelles chez des enfants de deux ans à deux ans et demi et chez des enfants de quatre ans a été testée par négativité de discordance. Si l'on applique aux formants des voyelles produits par un sujet adulte une transposition fréquentielle au niveau du retour acoustique, le sujet corrige immédiatement sa production de voyelles dans le sens opposé à la perturbation introduite. Chez l'enfant de deux ans à deux ans et demi, la production des formants F1 et F2 est plus instable et varie d'un énoncé à l'autre. Si on modifie le retour acoustique comme précédemment, il n'y a aucun effet sur la production vocale. Quelle explication peut-on proposer pour rendre compte de ces observations ? Elles sont inattendues parce qu'à cet âge, l'enfant a déjà acquis une bonne connaissance de sa langue maternelle. Il commence à balbutier avec une prosodie qui porte les caractéristiques de la langue à laquelle il est exposé. Pourquoi ne dirige-t-il pas de façon automatique sa production vocale comme le fait l'adulte ou l'enfant à partir de quatre ans ? Deux types de réponses sont avancées : soit le test lui-même ne convient pas pour un tout jeune enfant, soit le mécanisme de correction d'erreurs, fondé sur le retour acoustique de sa propre parole, ne peut se développer chez l'enfant que lorsqu'il a acquis une représentation interne robuste de sa propre expression vocale. Or, il y a de fait une immaturité phonologique chez l'enfant de cet âge. En raison de la descente du larynx, la valeur des formants des voyelles qu'il prononce change entre un et quatre ans. Par conséquent, les paroles qui vont servir de modèle à ses corrections vocales sont davantage celles des personnes de son entourage que les siennes. On sait aujourd'hui que ce ne sont pas les mêmes populations cellulaires qui sont activées ou inhibées lorsqu'un animal vocalise, et lorsqu'il écoute d'autres individus de l'espèce vocaliser. De même, chez l'homme, l'IRMf a montré que ce ne sont pas les mêmes régions qui sont activées ou dont l'activation est modifiée dans les conditions d'écoute de sa propre parole et de celle des autres. Il existe donc en quelque sorte deux mécanismes de rétrocontrôle auditif ou deux systèmes : l'un qui répond à la propre voix du locuteur et l'autre aux voix des autres locuteurs.

Quatrième cours. Les dysfonctionnements de la production vocale chez l'homme : une voie d'approche des mécanismes d'acquisition du langage

Le quatrième cours a porté sur l'origine évolutive des productions sonores à caractère social, et sur les atteintes de la production et de la perception de la voix, de la parole et du langage.

L'évolution des productions sonores chez les vertébrés nous a amenés à discuter le rôle de la vessie natatoire présente chez certains poissons. C'est un sac à paroi mince, rempli d'un gaz dont la composition est différente d'une espèce à l'autre, et que l'on trouve chez les poissons osseux. Elle les dispense de nager en permanence pour éviter de couler au fond de l'eau. Elle détermine la profondeur à laquelle le poisson flotte. Dans les profondeurs, la pression qui s'exerce sur les poissons est considérable. La densité du gaz de la vessie natatoire est multipliée par 200 à 2 km de profondeur et la flottaison est quasi-nulle. Lorsque le poisson remonte à la surface, le gaz de la vessie natatoire se dilate, son volume augmente, et il flotte. La vessie natatoire a cependant une autre fonction : c'est un organe acoustique. Elle est entourée, dans sa partie crânienne, de muscles dont la contraction à haute fréquence crée des vibrations, qui se propagent dans l'eau environnante. On détecte parmi ces émissions sonores des « cris » de détresse, de chasse ou d'appel à la reproduction. La contraction musculaire qui fait vibrer la vessie natatoire serait présente chez une centaine d'espèces parmi les 800 espèces de poissons téléostéens connues.

Chez les batraciens, nous nous sommes intéressés à *Xenopus laevis* (Yamaguchi & Kelley, 2000), (Yamaguchi *et al.*, 2008). Les vocalisations dans cette espèce sont de deux types : les unes, exclusivement produites par les mâles, correspondent à un appel sexuel, les autres, émises par les deux sexes, sont appelées appel « tic-tac ». Dans les deux cas, elles correspondent à des contractions des muscles du larynx, qui écartent brutalement une paire de disques créant un bref claquement. L'analyse comparative du développement des circuits vocaux indique que chez les larves ou les embryons de tous les vertébrés, le circuit neuronal vocal se forme à la charnière entre le rhombencéphale et la moelle épinière (Bass *et al.*, 2008).

Hormis les troubles de l'audition, qui se manifestent par l'élévation du seuil de la perception auditive, et qui affectent aussi la production de la parole en raison de la déficience de la boucle audio-phonatoire, d'autres atteintes plus spécifiques de la production et de la perception vocale ont été décrites (Graham & Fisher, 2013). Des anomalies de la perception de la voix sont rapportées dans le cadre des troubles autistiques avec une perte d'activation de la rive supérieure du STS en réponse à des stimulations vocales (Gervais *et al.*, 2004). Ces résultats ont été confirmés par d'autres travaux. La dyslexie développementale s'accompagne, quant à elle, d'une anomalie de perception des phonèmes. Les troubles de la production vocale sont nombreux : dysphonies, dyspraxies verbales développementales, apraxie de la parole chez l'enfant, déficits spécifiques du langage...

Plusieurs gènes ont été identifiés comme conférant une susceptibilité à une ou plusieurs de ces atteintes, mais les anomalies de ces gènes ne rendent compte que d'une toute petite fraction des personnes affectées par ces troubles.

Le cours s'est ensuite concentré sur l'atteinte du gène *FOXP2*. En 1990, est rapportée une grande famille atteinte d'un trouble héréditaire de la parole transmis sur le mode dominant (Hurst *et al.*, 1990). Ce déficit est d'abord décrit comme une apraxie verbale développementale. En 2002, ces patients sont réexaminés, et

l'absence de fluidité de la parole ainsi que les difficultés de la construction grammaticale des phrases sont soulignées. Les troubles sont alors considérés comme étant davantage de nature linguistique qu'en rapport avec une atteinte motrice oro-faciale (Fisher *et al.*, 1998). En 2001, le gène responsable est identifié (Lai *et al.*, 2001). Il code pour un facteur de transcription qui appartient à la famille des facteurs de transcription à domaine *forkhead*. Plus précisément, il appartient à la sous-famille FoxP, qui comporte quatre membres présents chez tous les vertébrés. La mutation trouvée substitue une histidine à une arginine (R553H) et se situe dans le domaine *forkhead*, domaine de liaison à l'ADN. Deux autres anomalies de ce gène seront rapportées chez des patients, dont l'une consiste en une translocation équilibrée, et l'autre en une mutation non-sens (R328X). La mise en évidence, l'année suivante (Enard *et al.*, 2002), d'une évolution particulièrement rapide de ce gène limitée à la lignée humaine va susciter un intérêt considérable, car elle est cohérente avec l'idée de l'implication de ce gène dans le langage. L'imagerie cérébrale chez les patients va révéler une réduction des noyaux caudés du striatum. D'un patient à l'autre, cette réduction paraît proportionnelle au degré de dyspraxie oro-faciale. La réduction d'autres régions cérébrales est observée, dont celle de l'aire de Broca. À l'inverse, la taille du gyrus temporal supérieur est accrue. L'imagerie fonctionnelle cérébrale met en évidence des défauts fonctionnels des aires motrices cérébrales. Elle révèle également d'autres anomalies de l'activité cérébrale. De fait, dans le cerveau humain, l'expression du gène *FOXP2* est très intense dans le striatum (noyau caudé et pallidum), mais elle l'est aussi dans l'aire qui lie l'aire de Broca et celle de Wernicke, le cortex frontal et occipital, le thalamus, le cervelet et le tronc cérébral. Pour comprendre le rôle de la protéine *FOXP2*, divers modèles animaux ont été générés ou sélectionnés : souris dont le gène orthologue est inactivé (mutation non-sens S321X), souris porteuses de la mutation R553H observée chez les patients (Groszer *et al.*, 2008), souris porteuses d'une version dite « humanisée » de ce gène par l'introduction de changements nucléotidiques qui convertissent la séquence de la protéine de souris en séquence humaine (Enard *et al.*, 2009), et enfin oiseaux chez lesquels la dégradation de l'ARNm codé par ce gène est obtenue par la technique d'interférence par ARN (Haesler *et al.*, 2007). L'ensemble des résultats indique que *FOXP2* intervient dans la production des vocalisations, qu'il est exprimé par les neurones épineux qui forment 90 % de la population cellulaire du striatum, qu'il joue un rôle critique dans l'accroissement de leurs dendrites et leur plasticité synaptique. L'évolution de la séquence nucléotidique du gène *FOXP2* continue d'être un élément important dans l'appréciation du rôle de ce gène dans le développement du langage. En effet, seuls cinq gènes ont été identifiés comme ayant subi une accélération évolutive dans la lignée humaine. Il s'agit de gènes impliqués dans la spermatogenèse ou le système immunitaire. *FOXP2* est le seul à être impliqué dans des fonctions cérébrales. L'évolution de la séquence du gène *FOXP2* a été étudiée dans des espèces qui apprennent les vocalisations : les baleines, l'éléphant africain et des dizaines d'espèces de chauves-souris, dont certaines sont douées d'apprentissages vocaux. Chez les chauves-souris et seulement chez elles, une évolution extrêmement rapide de la séquence de ce gène a été observée (Li *et al.*, 2007). Elle suggère que chez les chauves-souris, ce gène fait l'objet d'une sélection positive, comme dans la lignée humaine ; aucun des changements décelables d'acides aminés n'est identique à ceux observés dans la lignée humaine. L'idée a été avancée d'un rôle de l'évolution de *Foxp2* dans l'écholocation utilisée par les chauves-souris. Les

nouvelles techniques d'ingénierie génétique, qui permettent d'introduire plusieurs mutations dans un même gène et de modifier les génomes de n'importe quel animal, devraient contribuer à clarifier le rôle du gène *FOXP2*.

Références bibliographiques

Bass A.H., Gilland E.H., Baker R. (2008) « Evolutionary origins for social vocalization in a vertebrate hindbrain-spinal compartment », *Science*, 321, 417-21.

Behroozmand R., Korzyukov O., Larson C.R. (2011) « Effects of voice harmonic complexity on ERP responses to pitch-shifted auditory feedback », *Clin Neurophysiol*, 122, 2408-17.

Belin P., Zatorre R.C.J. (2003) « Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport* », 14, 2105-9.

Belin P., Zatorre R.J., Lafaille P., Ahad P., Pike B. (2000) « Voice-selective areas in human auditory cortex », *Nature*, 403, 309-12.

Bendor D., Wang X. (2005) « The neuronal representation of pitch in primate auditory cortex », *Nature*, 436, 1161-5.

Bestelmeyer P.E., Belin P., Grosbras M.H. (2011) « Right temporal TMS impairs voice detection », *Curr Biol*, 21, R838-9.

Blank H., Anwender A., von Kriegstein K. (2011) « Direct structural connections between voice- and face-recognition areas », *J Neurosci*, 31, 12906-15.

Blasi A., Mercure E., Lloyd-Fox S., Thomson A., Brammer M., Sauter D., Deeley Q., Barker G.J., Renval V., Deoni S., Gasston D., Williams S.C., Johnson M.H., Simmons A., Murphy D.G. (2011) « Early specialization for voice and emotion processing in the infant brain », *Curr Biol*, 21, 1220-4.

Brumm H., Zollinger S.A. (2011) « The evolution of the Lombard effect: 100 years of psychoacoustic research », *Behaviour*, 148, 1173-98.

Enard W., Gehre S., Hammerschmidt K., Holter S.M., Blass T., Somel M., Bruckner M.K., Schreiweis C., Winter C., Sohr R., Becker L., Wiebe V., Nickel B., Giger T., Muller U., Groszer M., Adler T., Aguilar A., Bolle I., Calzada-Wack J., Dalke C., Ehrhardt N., Favor J., Fuchs H., Gailus-Durner V., Hans W., Holzlwimmer G., Javaheri A., Kalaydjiev S., Kallnik M., Kling E., Kunder S., Mossbrugger I., Naton B., Racz I., Rathkolb B., Rozman J., Schrewe A., Busch D.H., Graw J., Ivandic B., Klingenspor M., Klopstock T., Ollert M., Quintanilla-Martinez L., Schulz H., Wolf E., Wurst W., Zimmer A., Fisher S.E., Morgenstern R., Arendt T., de Angelis M.H., Fischer J., Schwarz J., Paabo S. (2009) « A humanized version of *Foxp2* affects cortico-basal ganglia circuits in mice », *Cell*, 137, 961-71.

Enard W., Przeworski M., Fisher S.E., Lai C.S., Wiebe V., Kitano T., Monaco A.P., Paabo S. (2002) « Molecular evolution of *FOXP2*, a gene involved in speech and language », *Nature*, 418, 869-72.

Fant G. (1960) « Acoustic theory of speech production », Mouton, The Hague.

Fisher S.E., Vargha-Khadem F., Watkins K.E., Monaco A.P., Pembrey M.E. (1998) « Localisation of a gene implicated in a severe speech and language disorder », *Nat Genet* 18, 168-70.

Fitch W.T., Reby D. (2001) « The descended larynx is not uniquely human », *Proc Biol Sci*, 268, 1669-75.

Folland N.A., Butler B.E., Smith N.A., Trainor L.J. (2012) « Processing simultaneous auditory objects : infants' ability to detect mistuning in harmonic complexes », *J Acoust Soc Am*, 131, 993-7.

Formisano E., De Martino F., Bonte M., Goebel R. (2008) « “Who” is saying “what”? Brain-based decoding of human voice and speech », *Science*, 322, 970-3.

Gervais H., Belin P., Boddaert N., Leboyer M., Coez A., Sfaello I., Barthelemy C., Brunelle F., Samson Y., Zilbovicius M. (2004) « Abnormal cortical voice processing in autism », *Nat Neurosci*, 7, 801-2.

Goldstein M.H., King A.P., West M.J. (2003) « Social interaction shapes babbling : testing parallels between birdsong and speech », *Proc Natl Acad Sci U.S.A.*, 100, 8030-5.

Graham S.A., Fisher S.E. (2013) « Decoding the genetics of speech and language », *Curr Opin Neurobiol*, 23, 43-51.

Grossmann T., Oberecker R., Koch S.P., Friederici A.D. (2010) « The developmental origins of voice processing in the human brain », *Neuron*, 65, 852-8.

Groszer M., Keays D.A., Deacon R.M., de Bono J.P., Prasad-Mulcare S., Gaub S., Baum M.G., French C.A., Nicod J., Coventry J.A., Enard W., Fray M., Brown S.D., Nolan P.M., Paabo S., Channon K.M., Costa R.M., Eilers J., Ehret G., Rawlins J.N., Fisher S.E. (2008) « Impaired synaptic plasticity and motor learning in mice with a point mutation implicated in human speech deficits », *Curr Biol*, 18, 354-62.

Haesler S., Rochefort C., Georgi B., Licznarski P., Osten P., Scharff C. (2007) « Incomplete and inaccurate vocal imitation after knockdown of *FoxP2* in songbird basal ganglia nucleus Area X », *PLoS Biol*, 5, e321.

He C., Trainor L.J. (2009) « Finding the pitch of the missing fundamental in infants », *J Neurosci*, 29, 7718-8822.

Herbst C.T., Stoeger A.S., Frey R., Lohscheller J., Titze I.R., Gumpenberger M., Fitch W.T. (2012) « How low can you go ? Physical production mechanism of elephant infrasonic vocalizations », *Science*, 337, 595-9.

Hurst J.A., Baraitser M., Auger E., Graham F., Norell S. (1990) « An extended family with a dominantly inherited speech disorder », *Dev Med Child Neurol*, 32, 352-5.

Koda H., Nishimura T., Tokuda I.T., Oyakawa C., Nihonmatsu T., Masataka N. (2012) « Soprano singing in gibbons », *Am J Phys Anthropol*, 149, 347-55.

Kuhl P.K., Tsao F.M., Liu H.M. (2003) « Foreign-language experience in infancy : effects of short-term exposure and social interaction on phonetic learning », *Proc Natl Acad Sci USA*, 100, 9096-101.

Lai C.S., Fisher S.E., Hurst J.A., Vargha-Khadem F., Monaco A.P. (2001) « A forkhead-domain gene is mutated in a severe speech and language disorder », *Nature*, 413, 519-23.

Li G., Wang J., Rossiter S.J., Jones G., Zhang S. (2007) « Accelerated *FoxP2* evolution in echolocating bats », *PLoS One*, 2, e900.

Little A.C., Jones B.C., DeBruine L.M. (2011a) « Facial attractiveness : evolutionary based research », *Philos Trans R Soc Lond B Biol Sci*, 366, 1638-59.

Little A.C., Jones B.C., DeBruine L.M. (2011b) « The many faces of research on face perception », *Philos Trans R Soc Lond B Biol Sci*, 366, 1634-7.

MacDonald E.N., Johnson E.K., Forsythe J., Plante P., Munhall K.G. (2012) « Children’s development of self-regulation in speech production », *Curr Biol*, 22, 113-7.

Mampe B., Friederici A.D., Christophe A., Wermke K. (2009) « Newborns’ cry melody is shaped by their native language », *Curr Biol*, 19, 1994-7.

McGugin R.W., Gatenby J.C., Gore J.C., Gauthier I. (2012) « High-resolution imaging of expertise reveals reliable object selectivity in the fusiform face area related to perceptual performance », *Proc Natl Acad Sci USA*, 109, 17063-8.

Naatanen R., Gaillard A.W., Mantysalo S. (1978) « Early selective-attention effect on evoked potential reinterpreted », *Acta Psychol (Amst)* 42, 313-29.

Nishimura T., Mikami A., Suzuki J., Matsuzawa T. (2003) « Descent of the larynx in chimpanzee infants », *Proc Natl Acad Sci USA*, 100, 6930-3.

Oxenham A.J., Micheyl C., Keebler M.V., Loper A., Santurette S. (2011) « Pitch perception beyond the traditional existence region of pitch », *Proc Natl Acad Sci USA*, 108, 7629-34.

Pena M., Werker J.F., Dehaene-Lambertz G. (2012) « Earlier speech exposure does not accelerate speech acquisition », *J Neurosci*, 32, 11159-63.

Perani D., Saccuman M.C., Scifo P., Anwander A., Spada D., Baldoli C., Poloniato A., Lohmann G., Friederici A.D. (2011) Neural language networks at birth. *Proc Natl Acad Sci USA*, 108, 16056-61.

Perrachione T.K., Del Tufo S.N., Gabrieli J.D. (2011) « Human voice recognition depends on language ability », *Science*, 333, 595.

Rossion B. (2008) « La reconnaissance des visages », *Cerveau & Psycho*, 25, 58-65.

Stoeger A.S., Mietchen D., Oh S., de Silva S., Herbst C.T., Kwon S., Fitch W.T. (2012) « An asian elephant imitates human speech », *Curr Biol*, 22, 2144-8.

Striem-Amit E., Hertz U., Amedi A. (2011) « Extensive cochleotopic mapping of human auditory cortical fields obtained with phase-encoding fMRI », *PLoS One*, 6, e17832.

Thomas C., Avidan G., Humphreys K., Jung K.J., Gao F., Behrmann M. (2009) « Reduced structural connectivity in ventral visual cortex in congenital prosopagnosia », *Nat Neurosci*, 12, 29-31.

Tillmann B., Gosselin N., Bigand E., Peretz I. (2012) « Priming paradigm reveals harmonic structure processing in congenital amusia », *Cortex*, 48, 1073-8.

Titze I.R. (2000) « Principles of voice production », National Center for Voice and Speech, Iowa City, IA.

von Kriegstein K., Dogan O., Gruter M., Giraud A.L., Kell C.A., Gruter T., Kleinschmidt A., Kiebel S.J. (2008) « Simulation of talking faces in the human brain improves auditory speech recognition », *Proc Natl Acad Sci USA*, 105, 6747-52.

von Kriegstein K., Giraud A.L. (2006) « Implicit multisensory associations influence voice recognition », *PLoS Biol*, 4, e326.

Yamaguchi A., Gooler D., Herrold A., Patel S., Pong W.W. (2008) « Temperature-dependent regulation of vocal pattern generator », *J Neurophysiol*, 100, 3134-43.

Yamaguchi A., Kelley D.B. (2000) « Generating sexually differentiated vocal patterns : laryngeal nerve and EMG recordings from vocalizing male and female african clawed frogs (*Xenopus laevis*) », *J Neurosci*, 20, 1559-67.

Organisation des enseignements au Collège de France

Les cours portant sur « La voix : sa production, sa perception » se sont tenus les jeudis 13 décembre 2012, 10, 17 et 24 janvier 2013^a.

Ils ont été accompagnés des séminaires suivants^b :

– Nathalie Henrich, Laboratoire « Grenoble Images Parole Signal Automatique », (GIPSA), université Joseph Fourier, Grenoble : *Les mystères de la voix chantée, entre art et sciences* (13 décembre 2012) ;

– Shihab Shamma, University of Maryland, États-Unis & École normale supérieure, Paris : *Auditory representation of temporal information : basic mechanisms and percepts* (10 janvier 2013) ;

a. Les cours sont disponibles en vidéo sur le site Internet du Collège de France : <http://www.college-de-france.fr/site/christine-petit/course-2012-2013.htm> [Ndlr].

b. Les séminaires sont disponibles en audio sur le site Internet du Collège de France : <http://www.college-de-france.fr/site/christine-petit/seminar-2012-2013.htm> [Ndlr].

– Martine Hausberger, *Éthologie animale et humaine (ETHOS)*, Campus de Beaulieu, Rennes : *Le chant des oiseaux* (17 janvier 2013) ;

– Andrej Kral, Institute of Audioneurotechnology, Allemagne : *Cortical representation of temporal information : effects of auditory experience* (24 janvier 2013).

COURS À L'ÉTRANGER AU TITRE DU COLLÈGE DE FRANCE

Danemark

Hindsgavl, 29 octobre 2012 : « From deafness genes to sound processing in the cochlea ».

Sandjberg, 5 mai 2013 : « Usher syndrome : the gateway to finding the common mechanisms operating in visual and auditory sensory cells ».

ACTIVITÉ DE RECHERCHE DU LABORATOIRE

Les avancées les plus significatives réalisées au cours de l'année 2012-2013 ont porté sur :

1) l'identification de deux nouveaux gènes de surdités de transmission récessive : *TSPEAR*, responsable d'une surdité profonde (DFNB98) (Delmaghani *et al.*, 2012, *Hum Mol Genet*), et *OTOGL*, responsable d'une surdité modérée (DFNB84) (Bonnet *et al.*, 2013, *Gene*) ;

2) la caractérisation fine, au pôle apical des cellules sensorielles auditives, du déplacement du kinocil au cours de la différenciation de la touffe ciliaire, chez la souris : mise en évidence d'un mouvement brownien confiné des deux centrioles associés au kinocilium et d'un déplacement graduel (Lepelletier *et al.*, 2013, *Biophys J*) ;

3) les principes de stimulation des cellules sensorielles auditives par la membrane tectoriale (Lukashkin *et al.*, 2012, *Proc Natl Acad Sci USA*) ;

4) la mise en évidence, dans les processus caliciels des cellules photo-réceptrices de la rétine chez les primates, de la présence du complexe formé par cinq protéines impliquées dans le syndrome de Usher de type 1 : une avancée dans la compréhension de la pathogénie de cette rétinopathie (Sahly *et al.*, 2012, *J Cell Biol*) ;

5) la caractérisation de la spectrine $\beta 5$, une isoforme géante de la spectrine, comme partenaire de liaison de plusieurs protéines impliquées dans le syndrome de Usher de type 1, et d'un rôle possible dans le transport de l'opsine vers le segment externe des cellules photo-réceptrices (Papal *et al.*, 2013, *Hum Mol Genet*).

Identification de deux nouveaux gènes de surdités de transmission récessive : *TSPEAR*, responsable d'une surdité profonde (DFNB98), et *OTOGL*, responsable d'une surdité modérée (DFNB84)

(Delmaghani S., Aghaie A., Michalski N., Bonnet C., Weil D., Louha M., Loundon N., Verpy E., Smaghe L., Hardelin J.-P., Rouillon I., Jonard L., Couderc R., Gherbi S., Garabedian E.N., Denoyelle F., Marlin S., Petit C.)

En analysant une famille iranienne consanguine dans laquelle ségrège une surdité congénitale profonde, nous avons caractérisé un nouveau locus de surdité récessive,

DFNB98, sur le bras long du chromosome 21, dans un intervalle de 4,8 Mb. Par la technique de séquençage total des exons du génome, nous avons identifié une mutation tronquante du gène *TSPEAR*, présente à l'état homozygote chez les individus atteints. Ce gène code une protéine sécrétée, contenant sept domaines EAR (*epilepsy-associated repeats*). Nous l'avons détectée, par immunofluorescence, à la surface de la touffe ciliaire des cellules sensorielles auditives chez la souris. La famille des protéines EAR comporte six membres, dont quatre (LG1, LG2, VLG1, TSPEAR) sont impliqués dans des maladies concernant le système auditif (épilepsies avec symptômes auditifs ou surdités). Ainsi, cette famille de protéines semble jouer un rôle essentiel dans le développement et le fonctionnement du système auditif.

L'autre gène, *OTOGL*, a également été identifié par la technique de séquençage total des exons du génome, chez un enfant atteint d'une surdité modérée. Il code une protéine similaire à l'otogéline (*otogelin-like*), une protéine des membranes acellulaires de l'oreille interne. La protéine OTOGL est donc vraisemblablement impliquée elle aussi dans les propriétés mécaniques de la membrane tectoriale, qui joue un rôle essentiel dans la transmission du stimulus sonore aux cellules ciliées externes de la cochlée.

Mise en évidence du caractère graduel de la migration du kinocil dans les cellules ciliées cochléaires en cours de différenciation

(Lepelletier L., Boutet de Monvel J., Buisson J., Desdouets C., Petit C.)

La morphogénèse de la touffe ciliaire se produit chez la souris en concomitance avec l'établissement de la polarité planaire du neuro-épithélium cochléaire, durant sa période d'extension convergente. La polarisation de la touffe ciliaire dépend d'un cil primaire spécialisé, le kinocil, qui migre de façon dirigée du centre à la périphérie de la surface apicale de la cellule ciliée. Cette migration n'ayant pas été observée directement, on sait très peu de choses sur sa dynamique et sur les forces qui la dirigent. En exploitant le gradient de différenciation des cellules ciliées de la base à l'apex de la cochlée, nous avons d'abord reconstruit, par immuno-marquage et segmentation de contours cellulaires, une carte des étapes de la migration du kinocil des CCI à un stade donné le long de la cochlée. Par ailleurs, nous avons développé une approche de vidéo-microscopie *ex vivo* permettant de suivre, dans des explants de cochlée de souris en culture, les mouvements des centrioles des CCI rendus fluorescents par l'expression de centrine fusionnée à la GFP. Un algorithme de traque permet d'analyser ces mouvements avec une précision de l'ordre de 10 nm. À tous les stades examinés au cours de la migration du kinocil, nous avons trouvé que le centriole mère (ou corps basal) et le centriole fille associé restent confinés à la surface apicale des CCI, chacun effectuant un mouvement Brownien couvrant une région de 0,2-0,35 μm de diamètre. L'analyse de ce mouvement indique qu'il est confiné par une force de rappel non-linéaire de l'ordre de 0,1 pN, vraisemblablement imposée aux centrioles par des éléments du cytosquelette. Mises ensemble, ces analyses statique et dynamique révèlent que la migration du kinocil est un processus graduel étalé sur $\sim 2,5$ jours, ayant lieu en parallèle avec le processus d'alignement des cellules ciliées en rangées ordonnées. L'analyse du mouvement des centrioles des cellules ciliées dans des conditions normales et pathologiques devrait permettre d'éclaircir les structures qui contraignent leur mouvement et celui du kinocil.

Principes de stimulation des cellules sensorielles auditives par la membrane tectoriale

(Lukashkin A.N., Legan P.K., Weddell T.D., Lukashkina V.A., Goodyear R.J., Welstead L., Petit C., Russell I.J., Richardson G.P.)

L'épithélium sensoriel de la cochlée, l'organe de Corti, est idéalement positionné pour la détection des vibrations, en réponse au son, de deux membranes acellulaires : la membrane basilaire, sur laquelle repose l'organe, et la membrane tectoriale, qui le recouvre. Un processus actif d'amplification, impliquant la production de forces mécaniques par les cellules ciliées externes, module l'interaction de ces deux membranes. Cette amplification cochléaire multiplie localement l'amplitude de vibration d'un facteur voisin de 1000, déterminant ainsi la sensibilité de l'oreille aux sons de faible intensité, ainsi que son pouvoir de résolution en fréquence. Deux hypothèses ont été proposées pour expliquer le rôle de la membrane tectoriale dans la stimulation des cellules ciliées externes. Elle pourrait agir comme un résonateur ou comme une masse inertielle. Nous avons pu trancher entre ces hypothèses grâce à l'étude de souris mutantes *Otoa*^{EGFP/EGFP} dépourvues d'otoancorine, une protéine présente exclusivement dans la membrane tectoriale. Chez ces souris, la membrane tectoriale est détachée de son point d'ancrage dans la région limbique (située du côté des cellules ciliées internes), mais elle reste couplée aux touffes ciliaires des cellules ciliées externes. Il y a disparition complète de la résonance de la membrane tectoriale observée chez les souris non-mutantes. Or ce couplage « tronqué » ne perturbe pas la stimulation des cellules ciliées externes, et on observe une amplification et une sélectivité fréquentielle des vibrations de la membrane basilaire pratiquement normales chez les souris mutantes. C'est donc l'inertie de la membrane tectoriale, et non sa résonance, qui joue un rôle majeur dans la stimulation des cellules ciliées externes, en déterminant le mode de déflexion de leurs touffes ciliaires avec une phase ajustée pour permettre l'amplification cochléaire. En revanche, la sensibilité de détection des cellules ciliées internes, mesurée par les potentiels d'action de leur fibres afférentes, est altérée chez les animaux mutants, ce qui établit le rôle essentiel du flux liquidien produit par les forces de cisaillement entre la membrane tectoriale et la surface apicale de l'organe de Corti dans la stimulation des touffes ciliaires de ces cellules.

Mise en évidence, dans les processus caliciels des cellules photo-réceptrices de la rétine chez les primates, de cinq protéines impliquées dans le syndrome de Usher de type 1

(Sahly I., Dufour E., Schietroma C., Michel V., Bahloul A., Perfettini I., Pepermans E., Estivalet A., Carette D., Aghaie A., Ebermann I., Lelli A., Iribarne M., Hardelin J.-P., Weil D., Sahel J.-A., El-Amraoui A., Petit C.)

La pathogénie de la dystrophie rétinienne du syndrome de Usher de type 1 (USH1) reste totalement incomprise car les souris mutantes pour les différents gènes responsables de ce syndrome, bien qu'elles en reproduisent le déficit auditif, n'ont cependant pas de dégénérescence de la rétine. Nous avons montré que dans les cellules photo-réceptrices (cônes et bâtonnets) du singe macaque, cinq protéines USH1 (myosine-7a, cadhérine-23, protocadhérine-15, harmonine, et SANS) sont colocalisées aux interfaces membranaires entre les segments interne et externe

d'une part, et entre les processus caliciels et le segment externe d'autre part. Ce profil d'expression est conservé chez les amphibiens et chez l'homme, mais pas chez la souris, dont les cellules photo-réceptrices sont pratiquement dépourvues de processus caliciels. Nos résultats suggèrent que les protéines impliquées dans le syndrome USH1 contribuent à une ceinture d'adhérence autour de la région basolatérale du segment externe des photo-récepteurs chez l'homme, et qu'un défaut de cette structure est vraisemblablement à l'origine de la dégénérescence rétinienne dans ce syndrome. Une vision holistique se construit donc autour des gènes *USH1*. Ils codent pour un complexe d'adhérence liant les microvillosités (stéréocils des cellules sensorielles de l'oreille ou processus caliciels des cellules photo-réceptrices) entre elles ou à un cil ou une structure d'actine. Reste à élucider si ce complexe est associé dans les photo-récepteurs à une fonction de mécano-transduction comme il l'est dans la touffe ciliaire des cellules sensorielles de l'oreille.

Caractérisation de la spectrine $\beta 5$, une isoforme géante de la spectrine, comme partenaire de liaison de plusieurs protéines impliquées dans le syndrome de Usher de type 1, et d'un rôle possible dans le transport de l'opsine vers le segment externe des cellules photo-réceptrices

(Papal S., Cortese M., Legendre K., Sorusch N., Dragavon J., Sahly I., Shorte S., Wolfrum U., Petit C., El-Amraoui A.)

La spectrine $\beta 5$ est une isoforme géante de la spectrine β , dont nous avons précédemment révélé la présence dans la paroi latérale des cellules ciliées externes de la cochlée. Nous avons montré que cette spectrine interagit avec trois protéines impliquées dans le syndrome de Usher de type 1, dont la myosine VIIa, une protéine motrice associée au cytosquelette d'actine, ainsi qu'avec différentes protéines de la photo-transduction, dont la rhodopsine, et avec certaines protéines motrices associées aux microtubules (kinésine II et complexe dynéine). Compte tenu de la distribution polarisée de la spectrine $\beta 5$ entre l'appareil de Golgi et la base du segment externe des cellules photo-réceptrices, nous avons proposé que cette spectrine, capable de former des homodimères *in vitro*, contribue au transport de l'opsine et d'autres protéines vers les disques du segment externe des photo-récepteurs, siège de la photo-transduction, en couplant ces protéines aux moteurs moléculaires associés aux microfilaments et aux microtubules.

PUBLICATIONS

2013

Avan P., Büki B., Petit C. (2013) « Auditory distortions : origins and functions », *Physiol Rev*, 93, 1563-619.

Bonnet C., Louha M., Loundon N., Michalski N., Verpy E., Smaghe L., Hardelin J.-P., Rouillon I., Jonard L., Couderc R., Gherbi S., Garabedian E.N., Denoyelle F., Petit C., Marlin S. (2013) « Biallelic nonsense mutations in the otogelin-like gene (*OTOGL*) in a child affected by mild to moderate hearing impairment », *Gene*, 527, 537-40.

Boulay A-C, del Castillo FJ, Giraudet F, Hamard G, Giaume C, Petit C, Avan P, Cohen-Salmon M (2013) « Hearing is normal without connexin30 », *J Neurosci*, 33, 430-4.

El-Amraoui A., Petit C. (2013) « Cadherin defects in inherited human diseases », In *Progress in Molecular Biology and Translational Science*, Conn PM (ed.), Elsevier. Vol. 116, pp. 361-84.

Lepelletier L., Boutet de Monvel J., Buisson J., Desdouets C., Petit C. (2013) « Auditory hair cell centrioles undergo confined brownian motion throughout the developmental migration of the kinocilium », *Biophys J*, 105, 48-58.

Papal S., Cortese M., Legendre K., Soroush N., Dragavon J., Sahly I., Shorte S., Wolfrum U., Petit C., El-Amraoui A. (2013) « The giant spectrin β V couples the molecular motors to phototransduction and Usher syndrome type I proteins along their trafficking route », *Hum Mol Genet*, 22, 3773-88.

2012

Bonnet C., El-Amraoui A. (2012) « Usher syndrome (sensorineural deafness and retinitis pigmentosa) : pathogenesis, molecular diagnosis and therapeutic approaches », *Curr Opin Neurol*, 25, 42-49.

Delmaghani S., Aghaïe A., Michalski N., Bonnet C., Weil D., Petit C. (2012) « Defect in the gene encoding the EAR/EPTP domain-containing protein TSPEAR causes DFNB98 profound deafness », *Hum Mol Genet*, 21, 3835-44.

Fakin A., Glavac D., Bonnet C., Petit C., Hawlina M. (2012) « Fundus autofluorescence and optical coherence tomography in relation to visual function in Usher syndrome type 1 and 2 », *Vision Res*, 75, 60-70.

Lukashkin A.N., Legan P.K., Weddell T.D., Lukashkina V.A., Goodyear R.J., Welstead L., Petit C., Russell I.J., Richardson G.P. (2012) « A mouse model for human deafness DFNB22 reveals that hearing impairment is due to a loss of inner hair cell stimulation », *Proc Natl Acad Sci USA*, 109, 19351-56.

Petit C., El-Amraoui A., Avan P. (2012) « Audition : hearing and deafness », In *Neuroscience in the 21st century : from basic to clinical*. Pfaff DW (ed.). Heidelberg : Springer Verlag, pp. 675-741.

Safieddine S., El-Amraoui A., Petit C. (2011) « The auditory hair cell ribbon synapse : from assembly to function. *Annu Rev Neurosci* 35, 509-28.

Sahly I., Dufour E., Schietroma C., Michel V., Bahloul A., Perfettini I., Pepermans E., Estivalet A., Carette D., Aghaie A., Ebermann I., Lelli A., Iribarne M., Hardelin J.-P., Weil D., Sahel J.-A., El-Amraoui A., Petit C. (2012) « Localization of Usher 1 proteins to the photoreceptor calyceal processes, which are absent from mice », *J Cell Biol*, 199, 381-99.

PROGRAMME EUROPÉEN

FP7 EEC TREATRUSH: « Fighting blindness of Usher syndrome : diagnosis, pathogenesis and retinal treatment » (2010-2014) – Coordinatrice scientifique (10 laboratoires : France, Allemagne, Grande-Bretagne, Italie, Pays-Bas, Suisse, États-Unis.