

# Perception de la parole et IRM : réalisation, évaluation et validation d'un système permettant une stimulation sonore de qualité en cours de séquence IRM

*Alain Ghio\**, *Bernard Teston\**, *Cheryl Frenck-Mestre\**, *Corine Astésano\*\**, *Daniele Schön\*\**,  
*Bruno Nazarian\*\*\**, *Muriel Roth\*\*\**, *Jean-Luc Anton\*\*\**

\* Laboratoire Parole et Langage, Université de Provence, 29 Av. R. Schuman, 13621 Aix-en-Provence, France  
Tel : +33 (0)4 42 95 36 30, Fax : +33 (0)4 42 59 50 96, Mél : alain.ghio@lpl.univ-aix.fr, www.lpl.univ-aix.fr

\*\* Institut de Neurosciences Physiologiques et Cognitives, CNRS, Marseille, France

\*\*\* Centre d'IRM Fonctionnelle Cérébrale, Hôpital de la Timone, Marseille, France

## ABSTRACT

This study describes the design and the assessment of a MRI-compatible sound production hardware. This system was developed to permit auditory studies with Magnetic Resonance Imaging (MRI) techniques. An important disadvantage caused by the MR imager is the acoustic noise generated during data acquisition, due to the fast gradient switching interacting with the main magnetic field. Several solutions were explored to reduce noise and to provide audio stimuli with a reasonable quality. The sound production system was first tested by instrumental methods (sound level, spectral analysis). Finally, perceptual tests consisting in intelligibility, semantic decision and prosodic judgement were achieved to validate the installation.

## 1. INTRODUCTION

Depuis une dizaine d'année, l'imagerie par résonance magnétique (IRM) s'affirme comme une technique incontournable dans l'étude de la parole et du langage. Elle peut être utilisée dans le but d'obtenir des images anatomiques pour l'analyse des mécanismes de production, ceci en remplacement de la cinéradioscopie X désormais prohibée dans la recherche scientifique. L'IRM reste cependant principalement employée pour les études d'imagerie cérébrale fonctionnelle (IRMf) permettant la mise en évidence des substrats neuronaux en relation avec les processus sensoriels, moteurs ou cognitifs impliqués dans des tâches de production et de perception de la parole. Si les résultats obtenus ont souvent une portée considérable, la technique d'IRM comporte toutefois deux inconvénients importants pour les études en sciences du langage : la faible résolution temporelle dans l'acquisition des images (de l'ordre de la seconde, voire de plusieurs secondes pour un volume complet) et le bruit acoustique généré par le dispositif lors de l'enregistrement d'images, empêchant ainsi toute stimulation sonore de qualité, particulièrement pour des acquisitions rapides de type echo planar (EPI) très bruyantes.

Cette étude présente la réalisation et l'évaluation d'un système permettant, au cours d'un bloc continu d'acquisition EPI, une stimulation sonore de qualité. Cette initiative se démarque des méthodes connues de sparse imaging où des périodes « silencieuses » interrompent l'acquisition des images afin d'y insérer les stimuli sonores [1]. Ce travail a été effectué au centre IRMf de l'Hôpital de la Timone à Marseille sur un dispositif Bruker à 3 Tesla.

## 2. POSITION DU PROBLEME

Le bruit acoustique généré par un dispositif d'IRM peut être très intense. Ce phénomène est dû principalement à la commutation rapide des bobines électriques générant les gradients de champ magnétique nécessaires à une acquisition multidimensionnelle de l'image. Nos mesures réalisées à l'aide d'un sonomètre Brüel & Kjaer de type 2238 Mediator ont permis de relever, au centre du tube IRM, des niveaux sonores compris entre 105 et 125 dB(A) durant une séquence EPI (l'une des plus classiques en IRMf). Ces mesures sont analogues à celles de Moer et al. [2], pour qui le bruit peut atteindre, pour un dispositif à 3 Tesla, un niveau continu de 105 à 133 dB avec des pointes à 140 dB. De ce fait, les sujets doivent impérativement se protéger grâce à des bouchons d'oreille et un casque anti-bruit. Enfin, un tel niveau de bruit rend inacceptable toute expérimentation nécessitant une stimulation sonore, en particulier dans le cadre d'études sur la perception de la parole.

Le deuxième obstacle empêchant une transmission électro-acoustique fidèle en milieu IRM provient de l'interdiction d'employer tout matériau ferrique dans l'environnement magnétique du dispositif. Cela exclut donc haut-parleurs, casques électrodynamiques et autres transducteurs acoustiques traditionnels.

Notre objectif a été, dans un premier temps, de réaliser un dispositif de diffusion sonore amagnétique. Une fois ce système opérationnel, nos efforts ont porté sur la réduction du bruit de l'IRM et l'augmentation de l'efficacité de la stimulation de façon à obtenir des conditions expérimentales acceptables. Une évaluation nous a finalement permis de valider toute l'installation.

## 3. UN CASQUE AUDIOPHONIQUE AMAGNETIQUE A CAPSULES PIEZOELECTRIQUES

L'emploi de matériaux ferromagnétiques est totalement exclu dans un environnement IRM, d'une part à cause des forces exercées sur ce type de matériau par un champ de 3 Tesla (60 000 fois le champ magnétique terrestre), mais aussi à cause des perturbations et des inhomogénéités locales engendrées par ces objets sur le champ constant, entraînant d'importants artéfacts. Une des caractéristiques d'un matériau piézoélectrique réside dans sa capacité à se déformer sous l'effet d'une tension électrique. Il existe ainsi des capsules piézoélectriques en forme de disque pouvant faire office de membrane vibrante capable de générer un son. La plupart sont fabriquées à base de

céramique avec des électrodes en cuivre, matériaux compatibles IRM. En plaçant ces capsules dans la coque d'un casque anti-bruit, nous obtenons un transducteur audio amagnétique.

Différents types de capsules ont été testés. Nous avons retenu les KPS-100 de Kingstate pour leur plus grande bande passante et leur meilleur rendement. Toutefois, par construction physique, ce type de transducteur ne permet pas de produire une qualité sonore très fidèle. En effet, une partie du signal basse fréquence (en dessous de 200 Hz) est coupée et la mauvaise linéarité spectrale entraîne une coloration du son initial. L'introduction en amont d'un égaliseur augmentant les basses, réduisant les résonances et relevant les anti-résonances permet de rattraper partiellement les non linéarités spectrales. Par contre, le problème du faible rendement acoustique reste quasi insoluble. En effet, une fois placé sur la tête d'un sujet protégé du bruit environnant par des bouchons d'oreille, le casque doit être poussé à un tel niveau sonore que les capsules piézoélectriques saturent rapidement.

L'utilisation de ce type de casque reste donc possible mais en dehors des séquences d'acquisition, c'est-à-dire pour la communication entre l'opérateur et le sujet ou encore, en acquisition fractionnée. Ce type de séquence dite sparse imaging consiste à envoyer un stimulus sonore de quelques secondes durant une période silencieuse et de commencer l'acquisition dans un second temps [1]. La latence de 4 à 7 secondes de la réponse hémodynamique permet ainsi d'enregistrer l'activité cérébrale en « différé ». On alterne ainsi stimulation sonore, acquisition, stimulation, acquisition... d'où le nom de sparse imaging. Bien que largement utilisé, ce type de séquence implique une faible durée des stimuli (quelques secondes). De plus, d'un point de vue de l'exploitation des données, il reste moins efficace que des séquences dans lesquelles stimulation et acquisition sont continues.

Dans le but d'obtenir un système permettant une stimulation sonore de qualité en acquisition continue, nous nous sommes orientés vers d'autres solutions. Le premier axe de progression a consisté à tapisser le tube de l'aimant d'une surface d'absorbant acoustique. Bien que de faible épaisseur pour éviter de restreindre trop l'espace occupé par le sujet, ce procédé nous a permis de réduire le bruit de 1 à 10 dB selon la nature des séquences. Le second axe de progression s'est porté sur la réalisation d'un autre système de production sonore.

#### **4. UN CASQUE AUDIOPHONIQUE A TRANSMISSION PNEUMATIQUE**

Si, dans le cas du casque à capsules piézoélectriques, nous avons rendu amagnétique tout le dispositif de production sonore, nous avons par opposition envisagé la solution de garder une transduction électro-acoustique classique (électro-magnétique) suivie d'une transmission sonore purement pneumatique. Le principe est le suivant :

1. la source sonore, placée dans la salle de contrôle, délivre le signal de stimulation.
2. le signal se propage par des lignes de transmission filtrées lors du passage de la salle de contrôle à la salle

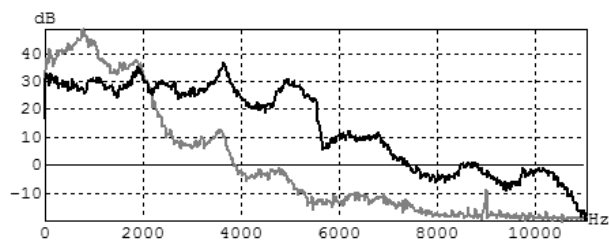
de l'aimant (cage de Faraday). Ce filtrage permet de limiter les interférences avec les RF induites par l'IRM.

3. le signal s'achemine à travers des câbles blindés et alimente un haut-parleur à chambre de compression de type Monacor KU-516 enfermé dans un coffret de blindage métallique en inox placé à proximité du lit du sujet. Ce type de transducteur électro-acoustique offre une grande pression sonore, une petite puissance d'ampli pour un son propre et bien audible. La bande passante est comprise entre 160 et 6500 Hz, valeurs acceptables pour de la transmission vocale.
4. un réducteur de forme conique permet une adaptation de diamètre de la sortie de la chambre de compression vers un tuyau en silicone Tygon de 7 mm ( $\varnothing$  intérieur).
5. le tuyau de 3 mètres de long transmet la vibration acoustique jusqu'au casque du sujet. Ce casque est un casque anti-bruit dont les coques sont alourdies par du silicone afin de renforcer l'absorption du bruit externe. Chaque coque est percée d'un trou ajusté pour permettre le passage étanche du tuyau.
6. à l'extrémité du tuyau est inséré un tube souple de 7 mm de diamètre extérieur qui termine la transmission à l'intérieur du conduit auditif. Une pâte en silicone durcissable permet d'une part d'assurer l'obturation du conduit auditif (atténuant ainsi le bruit de l'IRM), d'autre part de maintenir la position du tube terminal transmettant la stimulation sonore.

L'avantage de cette procédure réside dans la possibilité de fournir au niveau du tympan un signal potentiellement énergétique tout en atténuant fortement le bruit IRM grâce au casque anti-bruit et le bouchon d'oreille en silicone. Le rapport signal/bruit s'en trouve très fortement amélioré. En contrepartie, un risque demeure en cas de stimulation sonore excessive du fait de la proximité entre le tube terminal et le tympan. Un contrôle rigoureux du niveau est donc mis en place en limitant la puissance délivrée par l'amplificateur.

#### **5. EVALUATION INSTRUMENTALE**

Afin d'évaluer les caractéristiques de la transmission pneumatique, nous avons choisi comme source sonore un générateur de bruit blanc (Brüel & Kjaer modèle 1027) connecté à un amplificateur HI-FI Yamaha P2075. A l'autre extrémité de la transmission, nous avons connecté de façon étanche le tuyau de sortie au microphone d'un sonomètre Brüel & Kjaer de type 2238 Mediator. En plus de la mesure sonométrique, cet instrument joue aussi le rôle de préamplificateur qui, connecté à un système d'acquisition audio TASCAM US-122 lui-même branché à un ordinateur portable par le port USB, autorise l'enregistrement numérique du signal sonore dans la salle de l'aimant. La Figure 1 permet de mettre en évidence le filtrage passe-bas effectué par le tube de 3 mètres. Pour compenser ce phénomène qui se traduit perceptivement par un timbre sourd, nous avons introduit en amont de l'amplificateur un égaliseur réglé globalement pour atténuer les graves et rehausser les aigus (Figure 1).



**Figure 1** : Réponse spectrale de la stimulation pneumatique brute (en gris) et rectifiée (en noir)

Nous avons conscience des distorsions spectrales introduites par la ligne de stimulation sonore pneumatique, la réponse harmonique n'étant absolument pas plate. Toutefois, les conditions instrumentales en milieu IRM étant très fortement contraignantes, nous avons provisoirement décidé comme acceptable le dispositif. Nous ne perdons pas espoir de mieux linéariser la transmission en confectionnant sur mesure une étape de préfiltrage à la place de l'équaliseur.

## 6. EVALUATION PERCEPTIVE

### 6.1. Protocole expérimental

L'objectif de l'évaluation perceptive était d'estimer la possibilité de réaliser des expérimentations d'IRMf en continu avec stimulation auditive. Il s'agissait donc d'une validation préalable permettant d'envisager, par la suite, des expériences d'IRMf proprement dites. Nous avons procédé comme pour des tests comportementaux traditionnels mais dans l'environnement IRM. Nous n'avons absolument pas exploité les images fonctionnelles.

L'expérimentation se déroulait en trois phases. Une prise de contact et une courte séance d'entraînement à l'extérieur de l'aimant permettaient de vérifier la bonne compréhension des consignes. Le sujet était ensuite placé allongé dans le tube de l'aimant et disposait

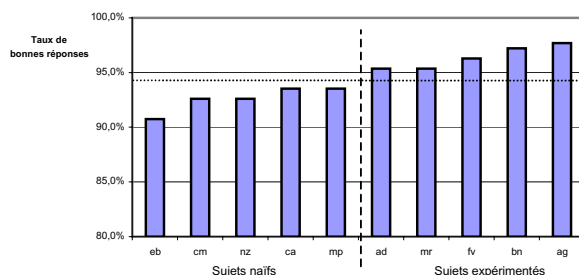
- du casque audio à transmission pneumatique.
- d'un coussin posé sous la nuque pour éviter la propagation des vibrations par conduction osseuse.
- d'un clavier de réponse.
- d'un miroir permettant au sujet de voir la projection sur un écran via un vidéo projecteur pour la lecture de consignes, tâches, propositions de réponses.

Une période préalable durant laquelle le sujet visualisait des images reposantes tout en écoutant de la musique relaxante permettait d'une part le réglage du volume sonore, d'autre part l'habitué aux conditions expérimentales peu banales (confinement, position couchée, vision par miroir...). Dans un dernier temps, durant la durée du test proprement dit, une séquence d'acquisition IRM EPI standard 123kHz, générant un niveau de bruit de 117 dB, était déclenchée en continu jusqu'à la fin du test. La stimulation audiovisuelle et le recueil des réponses étaient pilotés par un ordinateur portable placé dans la salle de contrôle à l'extérieur de l'aimant. Le pilotage expérimental était assuré par le logiciel PERCEVAL [3], un dispositif automatisé de tests de PERCEPTION et d'EVALUATION auditive et visuelle,

téléchargeable à l'adresse [www.lpl.univ-aix.fr/~lpldev/perceval](http://www.lpl.univ-aix.fr/~lpldev/perceval). L'intérêt de ce package est d'une part la possibilité de fabriquer des designs expérimentaux variés grâce à l'utilisation de fichiers scripts qui décrivent de façon flexible le déroulement du test. D'autre part, il offre la possibilité de brancher des boîtiers de réponse via le port USB, ce qui permet d'accepter un déport du sujet par rapport à l'ordinateur de pilotage. Dix sujets (5 naïfs, 5 expérimentés IRM) ont participé à l'évaluation perceptive qui a porté sur trois tests complémentaires.

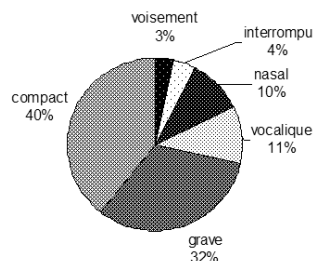
### 6.2. Test des paires minimales

Le test de paires minimales, comportant 216 stimuli, fournit une estimation quantitative d'un jugement d'intelligibilité [4]. Le sujet entend la production sonore d'un mot et doit choisir sa réponse parmi 2 mots en opposition sur un unique trait phonétique (ex : pile/bile, four/sourd, nonne/donne...). Sur les 2160 réponses recueillies tous sujets confondus, 2041 résultats étaient corrects, soit un taux moyen de 94.5 % d'intelligibilité (Figure 2). L'analyse des résultats des 5 sujets naïfs laisse apparaître un taux de 92.6 % alors que les sujets expérimentés (habituels à l'IRM) ont obtenu 96.4% d'intelligibilité, ce qui se rapproche des taux obtenus en condition normale de test. L'écart entre les deux groupes est significatif (T-test, p-value=0.0006). Ces analyses confirment, d'une part la bonne transmission audio, d'autre part l'intérêt d'habituer, voire d'entraîner les sujets aux conditions expérimentales de l'IRM.



**Figure 2** : Scores d'intelligibilité en fonction des sujets

De plus, sur les 119 erreurs relevées au total, 72 % concernent les traits acoustiques de compacité et d'acuité (Figure 3), qui correspondent, en terme articulatoire, au lieu d'articulation.



**Figure 3** : Typologie des erreurs d'intelligibilité

A titre d'exemple, on peut citer les oppositions de compacité b/g, d/g, p/k, t/k, z/ʒ, s/ʃ et celles d'acuité v/z, f/s, p/t, b/d, m/n, w/j. Là encore, nous retrouvons typiquement les traits les moins robustes de la communication parlée. Il faut enfin signaler que ces oppositions sont sensibles à la rectification spectrale

effectuée par l'équaliseur (Figure 1), un déplacement inadéquat des pôles d'énergie entraînant une catégorisation erronée.

## 6.2. Tests de décision sémantique et d'appréciation prosodique

Ce test, développé par Astésano et al. [5], est composé de 128 phrases se terminant par un mot trisyllabique. Le dernier mot de ces phrases est sémantiquement congruent (ex : « le concours a regroupé mille  candidats  ») ou incongru (ex : « le concours a regroupé mille  bigoudis  »). D'autre part, ce dernier mot peut être soit rythmiquement congruent (allongement final naturel), soit incongru (avec un allongement sur l'avant-dernière syllabe du dernier mot). Le corpus est donc constitué de 4 conditions expérimentales:

- nn) sémantique & rythme congruents
- np) sémantique congruente & rythme incongru
- sn) sémantique incongrue & rythme congruent
- sp) sémantique & rythme incongrus

Pour ce test, le corpus a été divisé en deux parties, chacune correspondant à une tâche différente. La première tâche consistait en une décision sémantique (« la phrase fait-elle sens ? » Oui/Non), qui correspond à un décodage global de la chaîne parlée. La deuxième partie du test consistait en une appréciation prosodique (« le rythme syllabique du dernier mot est-il correct ? » Oui/Non), qui correspond à un décodage local et fin au sein de la chaîne parlée. Dans chacune des tâches, les sujets étaient confrontés aux quatre conditions expérimentales décrites plus haut, nécessitant de sa part une attention sélective.

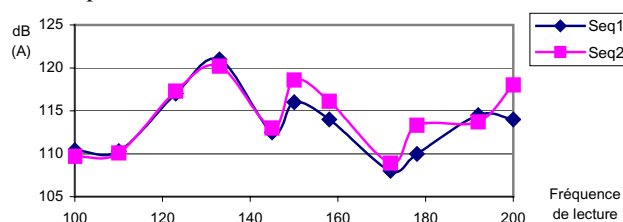
Pour la tâche de décision sémantique (59 stimuli x 10 sujets), les résultats laissent apparaître une moyenne de 95.9% de bonnes réponses (de 86.4% à 100% selon les sujets). Pour la tâche d'appréciation rythmique (60 stimuli x 10 sujets), les sujets ont obtenu en moyenne 94.8% de bonnes réponses (de 88.3% à 100%). L'analyse des erreurs met en évidence une baisse des réponses erronées dans les situations totalement congruentes ou incongrues (situations nn et sp) par rapport aux situations mixtes (sn et np), ce qui laisse penser plus à un problème de difficulté de réalisation de la tâche que de problèmes de stimuli mal transmis.

## 7. PERSPECTIVES

Le travail expérimental dans un environnement IRM reste très difficile et extrêmement contraignant pour tous les périphériques de stimulation ou de capteurs. En ce qui concerne la production sonore de qualité, il s'agit d'un engagement permanent pour gagner quelques décibels de signal et pour atténuer quelques décibels de bruit. Cette lutte doit s'opérer sur tous les fronts possibles.

Même si l'évaluation perceptuelle laisse apparaître des résultats très acceptables, nous avons conscience de la marge de manoeuvre qu'il nous reste encore à exploiter. L'évaluation instrumentale nous laisse entrevoir des améliorations potentielles, notamment en linéarisant mieux la réponse harmonique de la chaîne de transmission sonore. Le traitement acoustique de la salle de l'aimant par

des surfaces absorbantes devrait aussi permettre d'abaisser globalement le niveau général de bruit. De plus, un travail sur les sources elles-mêmes du bruit est en cours. En effet, nous étudions la programmation de séquences où la commutation des bobines de gradients (génératrices de bruit) pourrait être amortie grâce à des fronts de montée avachis. Dans la même veine, les modes de résonance propre du tube de l'aimant laissent apparaître des paramètres de programmation des séquences plus ou moins favorables en terme de bruit acoustique. Ainsi, selon la fréquence de lecture des points dans une séquence EPI, le bruit peut varier de 108 à 122 dBa (Figure 4). Bien évidemment, les caractéristiques générant un faible bruit sont à exploiter.



**Figure 4** : Niveau de bruit généré par des séquences EPI en fonction du paramétrage de la fréquence de lecture des points : mise en évidence des modes de résonance et anti-résonances acoustiques du tube de l'aimant.

## CONCLUSION

Notre objectif était de réaliser un système permettant une stimulation sonore de qualité au cours d'un bloc continu d'acquisition IRM. Cette initiative reste inédite en Europe à notre connaissance, les seuls dispositifs de production sonore existant ne permettant que des méthodes de sparse imaging. L'évaluation perceptuelle de la transmission audio nous permet de valider le dispositif. Les différentes pistes d'amélioration nous laissent entrevoir d'encore bien meilleurs résultats. De ce fait, cette réalisation technique ouvre d'excellentes perspectives aux études IRMf en perception de la parole et plus globalement auditive.

## BIBLIOGRAPHIE

- [1] P. Belin, R. Zatorre, R. Hoge, A. Evans, B. Pike, Event-Related fMRI of the Auditory Cortex, *NeuroImage*, volume 10, pages 417-429, 1999.
- [2] A. Moer, P. Pattynama, Acoustic Noise Concerns in Functional Magnetic Resonance Imaging, *Human Brain Mapping*, volume 20, pages 123-141, 2003.
- [3] C. André, A. Ghio., C. Cavé, B. Teston, PERCEVAL: a Computer-Driven System for Experimentation on Auditory and Visual Perception, *Proc. 15th ICPhS*, pages 1421-1424, Barcelone, 2003.
- [4] JP. Peckels, M. Rossi. A., Le test de diagnostic par paires minimales. Adaptation au français du Diagnostic Rhyme Test de Voiers, *Revue d'Acoustique*, volume 27, pages 245-262, 1973.
- [5] C. Astésano, C. Magne, M. Besson, Influence du rythme sur le traitement sémantique en français: approches comportementale et électrophysiologique. *Soumis dans ce volume*.