

## Demande de BQR Supmicrotech 2024

**Titre du projet :** DAVID v2

**Description rapide :** Validation et valorisation du dispositif de « feedback vocal » DAVID v2 sous LabVIEW/CompactRIO

**Nom et département du responsable du projet :** Patrick Nectoux (Ingénieur de recherche CNRS), Institut FEMTO-ST, Département AS2M.

**Durée du projet :** 2 ans

**Tableau de financement :**

	Investissement (TTC)	Fonctionnement (TTC)
Montant de la dotation demandée à SUPMICROTECH - ENSMM	8000€ (une plateforme CompactRIO + cartes d'E/S audio, devis ci-joint 7804€ + petite connectique)	3000€ (gratification pour stagiaire PFE Supmicrotech) 2000€ (frais de publication article) 2000€ (frais de mission pour démonstrateur)
Co-financement acquis (A) ou demandés (D)	-	-
Coût total du projet	8000	7000

## 1. Descriptif synthétique du projet

Le contrôle moteur de la parole est une problématique de recherche majeure en neurosciences ([Hickok, Nature Neuroscience Reviews 2012](#)). Cette fonction biologique est étudiée avec des dispositifs expérimentaux dit de « *altered auditory feedback* » (AAF) permettant de perturber le retour auditif des locuteur (ex. parler en entendant sa voix rendue artificiellement plus haute ou plus basse que ce qu'elle est en réalité) et de mesurer comment le locuteur s'y adapte ([Chang et al. PNAS 2013](#)). Or, les dispositifs actuellement utilisés pour faire ces transformations sont principalement des dispositifs commerciaux détournés de leur usage (ex. hardware de production musicale). Ces dispositifs manquent de flexibilité d'utilisation et leurs propriétés (ex. latence entrée-sortie) ne sont jamais validées.

Depuis 1 ans, le porteur et son équipe développe un dispositif AAF basé sur un contrôleur temps-réel NI CompactRIO (FPGA et processeur temps-réel), et qui a le potentiel de devenir le *gold standard* du domaine. Le dispositif est aujourd'hui fonctionnel, en ce qu'il permet en temps-réel la mesure de la hauteur de la parole produite et la transformation de la hauteur de la parole entendue en retour par un locuteur. Cependant, il manque plusieurs étapes permettant qu'il puisse être utilisé par des laboratoires de recherche extérieurs :

- d'une part, une validation expérimentale (campagne de mesure de latence et de validité de la transformation) ;
- d'autre part, un interfaçage de l'API du dispositif en python, pour pouvoir être utilisé dans les protocoles expérimentaux de psychologie/neuroscience, et sa documentation.

Ce sont ces deux dernières étapes (validation et valorisation), pour lesquelles le projet est actuellement sans ressources propres, qui font l'objet de cette demande BQR.

Cette demande est donc une véritable demande d'amorçage : avec la complétion de ces 2 étapes, il est attendu d'une part que le dispositif soit mis en test dans plusieurs laboratoires nationaux et internationaux (*LPNC, Université de Grenoble-Alpes ; Center for Music and the Brain, Université d'Aarhus, DK* ; sont déjà en attente), et d'autre part que ce dispositif fournisse la base d'un futur dépôt de projet collaboratif type ANR. Il s'agit également d'un projet de retour à la recherche du porteur, nouvellement promu ingénieur de recherche CNRS et qui a changé de problématique depuis le travail sur la plateforme PHM Pronostia ([Nectoux et al. IEEE PHM 2012](#) – 830 citations). Enfin, une version précédente du dispositif (DAVID v1), purement logicielle et développé à l'IRCAM ([Aucouturier et al. PNAS, 2016](#)), avait valu à l'établissement une belle réputation de recherche (3000 téléchargements entre 2018 et 2021). Il est donc attendu que ce nouveau dispositif apporte un beau rayonnement extérieur à Supmicrotech, bien aligné sur la nouvelle image de précision et de technologie pour la santé de la marque - en plus de fournir un nouveau support de manip pour des TP d'automatique et de traitement du signal aux étudiants de l'école.

## 2. État de l'art, objectifs, résultats attendus et planning prévisionnel

**État de l'art :** Le contrôle moteur de la parole est une problématique de recherche majeure en neurosciences (Hickok, 2012). Son étude a des implications à la fois fondamentales (car le contrôle articulé de la parole est considéré comme une propriété unique du cortex humain parmi les mammifères – Fitch et al. *Science Advances* 2016) et cliniques (car elle permet des avancées thérapeutiques et technologiques dans le domaine de prothèses vocales – Metzger et al. *Nature* 2023).

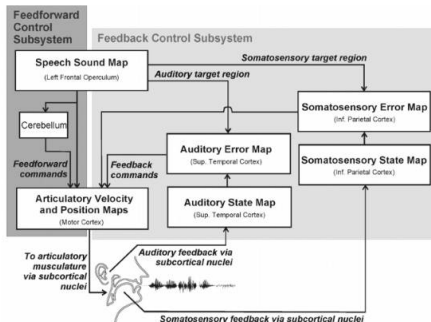


Figure 1: Mécanismes directs et de rétroaction dans la production vocale. Figure adaptée de Guenther (2007).

Comme tout dispositif de contrôle moteur biologique ou artificiel, la production de la parole implique un sous-système de contrôle « feed-forward », capable de traduire une cible auditive (par exemple un son /a/ à la fréquence de 150Hz) en une séquence motrice des structures du larynx et de la bouche permettant de produire le son correspondant, et un sous-système de « feedback » sensoriel, permettant de détecter d'éventuels erreurs entre la cible et le résultat obtenu et de les compenser en temps-réel (Guenther, 2007 ; Figure 1). Ce système boucle-fermée biologique est étudié expérimentalement en utilisant un 2<sup>e</sup> système boucle-fermée, artificiel celui-ci, qui est utilisé pour perturber le retour auditif des locuteurs et mesurer comment le locuteur s'y adapte (Figure 2 – gauche ; Chang et al. *PNAS* 2013).

Dans une expérience typique, on demande au participant de vocaliser un son stable avec une hauteur maîtrisée (ex. un /a/ avec une fréquence fondamentale de 150Hz), dans un microphone et en s'écoutant au casque ; à son insu, on introduit dans son retour audio une transformation de la hauteur du son, donnant l'impression au participant que le son produit est plus bas que ce qu'il pensait produire ; on observe alors une compensation automatique, et souvent inconsciente, dans la production du son, visant par exemple à en augmenter la fréquence fondamentale pour compenser le changement entendu, perturbation dont on peut étudier la dynamique (latence, amplitude, etc.) pour caractériser le système biologique (Figure 2 – droite ; Burnett & Larson, 2002).

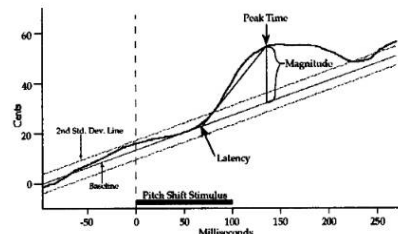
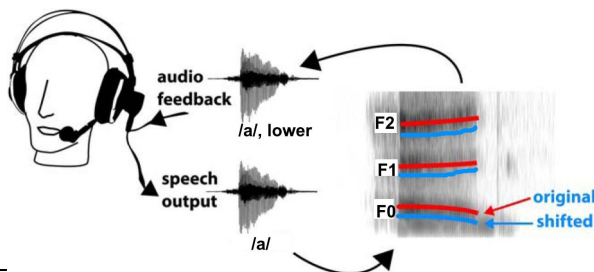


Figure 2 –  
Gauche :

Dispositif de feedback vocal altéré utilisé expérimentalement pour étudier la production vocale. Figure adaptée de Parrell et al. 2017. Droite : Résultat d'une expérience typique de perturbation de hauteur négative (« pitch shift stimulus ») appliquée pendant la production d'un son, et résultant en une compensation positive de la fréquence fondamentale produite 100ms plus tard. Figure adaptée de Burnett & Larson, 2002.

Les dispositifs utilisés pour faire ces transformations sont principalement des dispositifs commerciaux détournés de leur usage (ex. hardware de production musicale, comme l'*Ultra-harmonizer* de Eventide dans Burnett & Larson, 2002) ou des dispositifs logiciels maison (logiciel Matlab *Audapter*, Tourville et al. 2013). Or les propriétés temps-réel de ces dispositifs sont mal spécifiées et très rarement validées. Par exemple, un dispositif logiciel comme *Audapter* dépend des capacités d'entrée-sortie audio d'un système d'exploitation non temps-réel, typiquement Windows. Ceci entraîne, d'une part, des latences fixes de plusieurs 10ms qui perturbent le phénomène étudié et contraignent l'expérimentation à la seule observation de sons tenus et non de parole continue (Stuart et al. 2002) ; d'autre part, des variations de latence incontrôlables de l'ordre de 10-100ms qui viennent brouter toute la démarche expérimentale. Il est à noter au passage qu'à cause de ces limitations, presque rien n'est connu aujourd'hui sur les « constantes de temps » de la boucle de production de la parole humaine, et que l'impact de latences inférieures à 50ms est peu ou pas connu (Stuart et al. 2002) – car ces latences ne sont pas réalisables par les dispositifs actuels.

**Objectifs :** Depuis 1 ans, le porteur et son équipe développe un dispositif temps-réel basé sur un contrôleur temps-réel LabVIEW/CompactRIO (FPGA et processeur temps-réel) dédié au feedback vocal altéré, et qui a le potentiel technologique de devenir le *gold standard* du domaine. Celui-ci résout notamment le problème de la latence d'acquisition (réduite à 1ms avec des cartes d'entrée-sortie audio dédiée et un traitement en FPGA) et de sa maîtrise en cours de traitement (les transformations et leur contrôle étant implémentées sur processeur temps-réel dédié, à latence fixe). Le dispositif est aujourd'hui fonctionnel (stage PFE 2022 de Paul Maublanc), mais il manque deux étapes permettant qu'il puisse être utilisé de façon expérimentale par des laboratoires extérieurs, et qui font l'objet de cette demande BQR.

- Objectif 1 : validation expérimentale des capacités du dispositif

- campagne de mesure de la latence entrée/sortie du dispositif en utilisation, comparée avec le dispositif état de l'art *Audapter*.
- étude expérimentale de l'impact de la latence dans la bande (0-50ms) sur la boucle de production de la parole (étude expérimentale sur N=10 volontaires sains, selon un protocole psychophysique de *just noticeable difference* JND)
- validation de la transformation de hauteur, avec analyse acoustique des différences entre voix produite et voix entendue.

- Objectif 2 : interfaçage du dispositif en python et documentation

Afin de pouvoir être utilisé dans les protocoles expérimentaux de psychologie/neuroscience, le dispositif doit pouvoir être piloté à distance par un script expérimental, typiquement développé en python. Ce type d'interface (API) permet, par exemple, de présenter à un participant une série d'essais automatiques avec des transformations de hauteurs et/ou de latences variables d'essai et essai, sans nécessiter d'intervention manuelle sur le GUI du dispositif. Cette interface de programmation doit d'autre part être documentée de façon à être utilisable par de non-spécialistes.

**Résultats attendus :** Avec la complétion de ces 2 objectifs, il est attendu

1. Un démonstrateur du dispositif, mis en test dans plusieurs laboratoires nationaux et internationaux pour la réalisation d'expériences de feedback vocal altéré, et donnant lieu à des collaborations de recherche et des co-rédactions d'article.  
→ Nous avons déjà recueilli plusieurs déclarations d'intérêt de collègues français (Equipe de Louise Goupil, LPNC, Université de Grenoble-Alpes) et internationaux (équipes de Boris Kleber, Université d'Aarhus, DK ; de Ana Pinheiro, Université de Lisbonne, Portugal).
2. La rédaction d'un article de revue de méthodologie (target : Nature Methods, Behaviour Research Methods, Neuroscience Methods) décrivant l'implémentation du dispositif et sa validation expérimentale  
→ Deux précédents articles, décrivant une version précédente, purement logicielle, du dispositif (DAVID v1) publiés dans PNAS 2018 et BRM 2019 ont été cités resp. 75 et 45 fois.
3. La mise à disposition du logiciel sous forme de logiciel libre, open-source, contribuant au rayonnement de l'établissement dans les cercles de recherche, bien aligné sur la nouvelle image de précision et de technologie pour la santé de la marque et aux directives des tutelles, comme déjà réalisé pour plusieurs de nos résultats précédents ([github.com/neuro-team-femto](https://github.com/neuro-team-femto)).  
→ une version précédente du dispositif (DAVID v1), purement logicielle et développée à l'IRCAM (<https://forum.ircam.fr/projects/detail/david/>), avait valu à l'établissement une belle réputation de recherche (3000 téléchargements entre 2018 et 2021).
4. L'utilisation du démonstrateur comme nouveau support de manip pour des TP d'automatique et de traitement du signal des étudiants de l'école.  
→ participation de personnel ENSMM (Prof. Yongxin Wu) au projet, pour cadrage automatique et pédagogique

**Planning prévisionnel :**

- Février 2024 : dépense d'investissement (achat nouvelle génération CompactRIO + cartes d'acquisition)
- Mars – Août 2024 : projet PFE d'étudiant Supmicrotech (profil instrumentation/temps-réel), réalisation des campagnes de mesure et implémentation de l'API
- Septembre – Décembre 2024 : Rédaction d'un article et mise à disposition du code en open-source
- Printemps 2025 : Missions de démonstration du dispositif dans les laboratoires partenaires
- Septembre 2025 : Intégration du démonstrateur dans les cours de traitement du signal et d'automatique de l'établissement.

### 3. Expertise des participants dans le domaine considéré

Le porteur, **Patrick Nectoux**, ingénieur de recherche CNRS, est spécialiste d'instrumentation au département AS2M, FEMTO-ST. Jusqu'à récemment responsable du service commun électronique et instrumentation (SCEI) de l'Institut, il est impliqué en interne et en externe dans de nombreux projets de développement d'instrumentation temps-réel, ainsi que formateur LabView auprès de l'école doctorale ED SPIM. Depuis Jan. 2023, Patrick Nectoux est coordinateur national du Réseau des Electroniciens du CNRS. La demande BQR est un projet de retour à la recherche pour le porteur, nouvellement promu ingénieur de recherche et qui a changé de problématique depuis son travail sur la plateforme PHM Pronostia ([Nectoux et al. IEEE PHM 2012](#) – 830 citations).

Nectoux, P., Gouriveau, R., Medjaher, K., Ramasso, E., Chebel-Morello, B., Zerhouni, N., & Varnier, C. (2012, June). PRONOSTIA: An experimental platform for bearings accelerated degradation tests. In *IEEE International Conference on Prognostics and Health Management, PHM'12*. (pp. 1-8). IEEE Catalog Number: CPF12PHM-CDR.

Javed, K., Gouriveau, R., Zerhouni, N., & Nectoux, P. (2014). Enabling health monitoring approach based on vibration data for accurate prognostics. *IEEE Transactions on industrial electronics*, 62(1), 647-656.

**Yongxin Wu**, personnel Supmicrotech, est maître de conférences en automatique à Supmicrotech-ENSM ou il enseigne notamment les modules de bases de l'automatique ainsi que les TP. Titulaire d'un master et d'une thèse d'Automatique de l'Université Claude Bernard Lyon 1, ses travaux de recherche au département AS2M, FEMTO-ST concernent les systèmes Hamiltoniens à port, et les systèmes à paramètres distribués. Yongxin Wu est également co-directeur du Master International on Control for Green Mechatronics (Greem). Sa contribution au projet consiste au cadrage pédagogique de la maquette du dispositif, destinée aux travaux pratiques d'automatique de l'établissement, et il contribuera également à la conception du contrôleur boucle-fermée du dispositif.

Mattioni, A., Wu, Y., Le Gorrec, Y., & Zwart, H. (2022). Stabilization of a class of mixed ODE–PDE port-Hamiltonian systems with strong dissipation feedback. *Automatica*, 142, 110284.

Toledo, J., Ramirez, H., Wu, Y., & Gorrec, Y. L. (2023). Linear Matrix Inequality design of observer-based controllers for port-Hamiltonian systems. *IEEE Transactions on Automatic Control*

**Jean-Julien Aucouturier**, directeur de recherche CNRS au département AS2M, FEMTO-ST, est spécialiste de traitement du signal vocal et de neurosciences cognitives. Concepteur d'une précédente version, purement logicielle, du dispositif, JJA amène au projet une expertise sur la validation expérimentale de dispositifs de vocal feedback, sur l'implémentation d'algorithmes de transformation et analyse de voix, ainsi qu'une connaissance des enjeux fondamentaux dans le domaine des neurosciences de la voix.

Aucouturier, J. J., Johansson, P., Hall, L., Segnini, R., Mercadié, L., & Watanabe, K. (2016). Covert digital manipulation of vocal emotion alter speakers' emotional states in a congruent direction. *Proceedings of the National Academy of Sciences*, 113(4), 948-953.

Rachman, L., Liuni, M., Arias, P., Lind, A., Johansson, P., Hall, L., ... & Aucouturier, J. J. (2018). DAVID: An open-source platform for real-time transformation of infra-segmental emotional cues in running speech. *Behavior research methods*, 50(1), 323-343.

## 4. Références

- Burnett, T. A., & Larson, C. R. (2002). Early pitch-shift response is active in both steady and dynamic voice pitch control. *The Journal of the Acoustical Society of America*, 112(3), 1058-1063.
- Chang, E. F., Niziolek, C. A., Knight, R. T., Nagarajan, S. S., & Houde, J. F. (2013). Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proceedings of the National Academy of Sciences*, 110(7), 2653-2658.
- Fitch, W. T., De Boer, B., Mathur, N., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science advances*, 2(12), e1600723.
- Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of communication disorders*, 39(5), 350-365.
- Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature reviews neuroscience*, 13(2), 135-145.
- Metzger, S. L., Littlejohn, K. T., Silva, A. B., Moses, D. A., Seaton, M. P., Wang, R., ... & Chang, E. F. (2023). A high-performance neuroprosthesis for speech decoding and avatar control. *Nature*, 1-10.
- Parrell, B., Agnew, Z., Nagarajan, S., Houde, J., & Ivry, R. B. (2017). Impaired feedforward control and enhanced feedback control of speech in patients with cerebellar degeneration. *Journal of Neuroscience*, 37(38), 9249-9258.
- Stuart, A., Kalinowski, J., Rastatter, M. P., & Lynch, K. (2002). Effect of delayed auditory feedback on normal speakers at two speech rates. *The Journal of the Acoustical Society of America*, 111(5), 2237-2241.
- Tourville, J. A., Cai, S., & Guenther, F. (2013, June). Exploring auditory-motor interactions in normal and disordered speech. In *Proceedings of Meetings on Acoustics* (Vol. 19, No. 1). AIP Publishing.




ENSMM  
M. NECTOUX Patrick  
24 rue Alain Savary  
25000 BESANÇON  
FRANCE

Date de l'offre: 08/09/2023  
Offre valable jusqu'au: 08/10/2023  
Tel: +33 (03) 81 40 27 86  
Fax: (3) 81 88 57 14  
Contact No: 523989

### Offre de prix N° 2663505 Version 1

Nous avons le plaisir de vous adresser la présente offre de prix. Nous vous invitons à préciser le numéro de cette offre sur votre bon de commande.

Opt.	Référence	Description	Qté	PU net HT	Remise %	Montant net HT
		NI CRIO ID de configuration: <a href="#">CR6613670</a> 				
1.1	<a href="#">786425-01</a>	cRIO-9054, Double cœur 1,33 GHz, FPGA Artix-7 100T, 4 emplacements, RT, non-XT  Contacter le service client pour avoir une information de disponibilité <i>Pays d'Origine: Hungary</i>	1	2,980.00	10.00%	2,682.00
1.2	<a href="#">SRV-CR6613670</a>	Programme SSP (Standard Service Program) pour les systèmes CompactRIO pendant 3 ans  Contacter le service client pour avoir une information de disponibilité <i>Pays d'Origine: Hungary</i>				329.69
1.3	<a href="#">783814-01</a>	NI 9251 Module d'entrée analogique différentiel 2 voies, mXLR, 3 Veff, 24 bits, 102,4 kéch./s/voie  Contacter le service client pour avoir une information de disponibilité <i>Pays d'Origine: Hungary</i>	1	1,920.00	10.00%	1,728.00
1.4	<a href="#">783467-01</a>	NI 9260 Module sortie analogique Mini XLR, +/- 4,24 V, 24 bits, 51,2 kéch./s/voie, 2 voies  Contacter le service client pour avoir une information de disponibilité <i>Pays d'Origine: Hungary</i>	1	1,960.00	10.00%	1,764.00
Total HT:						EUR 6,503.69
TVA:						EUR 1,300.74
Total TTC:						EUR 7,804.43

#### Offre de prix établie en: Euro

#### Information complémentaire:

- Conditions de paiement: 45 jours fin de mois
- Ceci est un devis confidentiel de NI

Pour les commandes contenant des logiciels en souscription, ces logiciels seront automatiquement renouvelés à la fin de leur période de validité aux prix en vigueur à la date du renouvellement, à moins que d'autres conditions ne s'appliquent sur la base d'un accord écrit entre les