



Synchronisation vocale et mouvement compensé en reconstruction pour une ciné IRM de la parole

Pierre-André Vuissoz, Freddy Odille, Emmanuel Vincent, Jacques Felblinger,
Yves Laprie

► To cite this version:

Pierre-André Vuissoz, Freddy Odille, Emmanuel Vincent, Jacques Felblinger, Yves Laprie. Synchronisation vocale et mouvement compensé en reconstruction pour une ciné IRM de la parole. 2e Congrès de la SFRMBM (Société Française de Résonance Magnétique en Biologie et Médecine), Mar 2015, Grenoble, France. hal-01104230

HAL Id: hal-01104230

<https://inria.hal.science/hal-01104230>

Submitted on 6 Oct 2015

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

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

Pierre-André Vuissoz^{1,2}, Freddy Odille^{1,2}, Emmanuel Vincent^{3,4}, Jacques Felblinger^{5,6}, Yves Laprie^{3,7}

1. Imagerie Adaptative Diagnostique et Interventionnelle, Université de Lorraine, Nancy, France **2.** U947, INSERM, Nancy, France **3.** LORIA, Université de Lorraine, Nancy, France **4.** Inria, Nancy, France **5.** CHU Nancy, Nancy, France **6.** CIC-IT 1433, INSERM, Nancy, France **7.** UMR 7503, CNRS, Nancy, France

But de l'étude

L'imagerie dynamique du conduit vocal permet d'étudier et de modéliser la production de la parole.

Travaux antérieurs :

- La fluoroscopie à rayon X : rayon ionisant;
- L'IRM temps réel [1] reste limitée en résolution;
- La synchronisation de l'IRM par un système acoustique est possible [2] mais nécessite une parfaite reproductibilité de la parole.

Dans ce travail :

- Une acquisition IRM non synchronisée avec un enregistrement de la parole utilisant un microphone compatible;
- Une reconstruction Cinématique compensée en mouvement.



Matériel et méthodes

➤ Acquisition des données

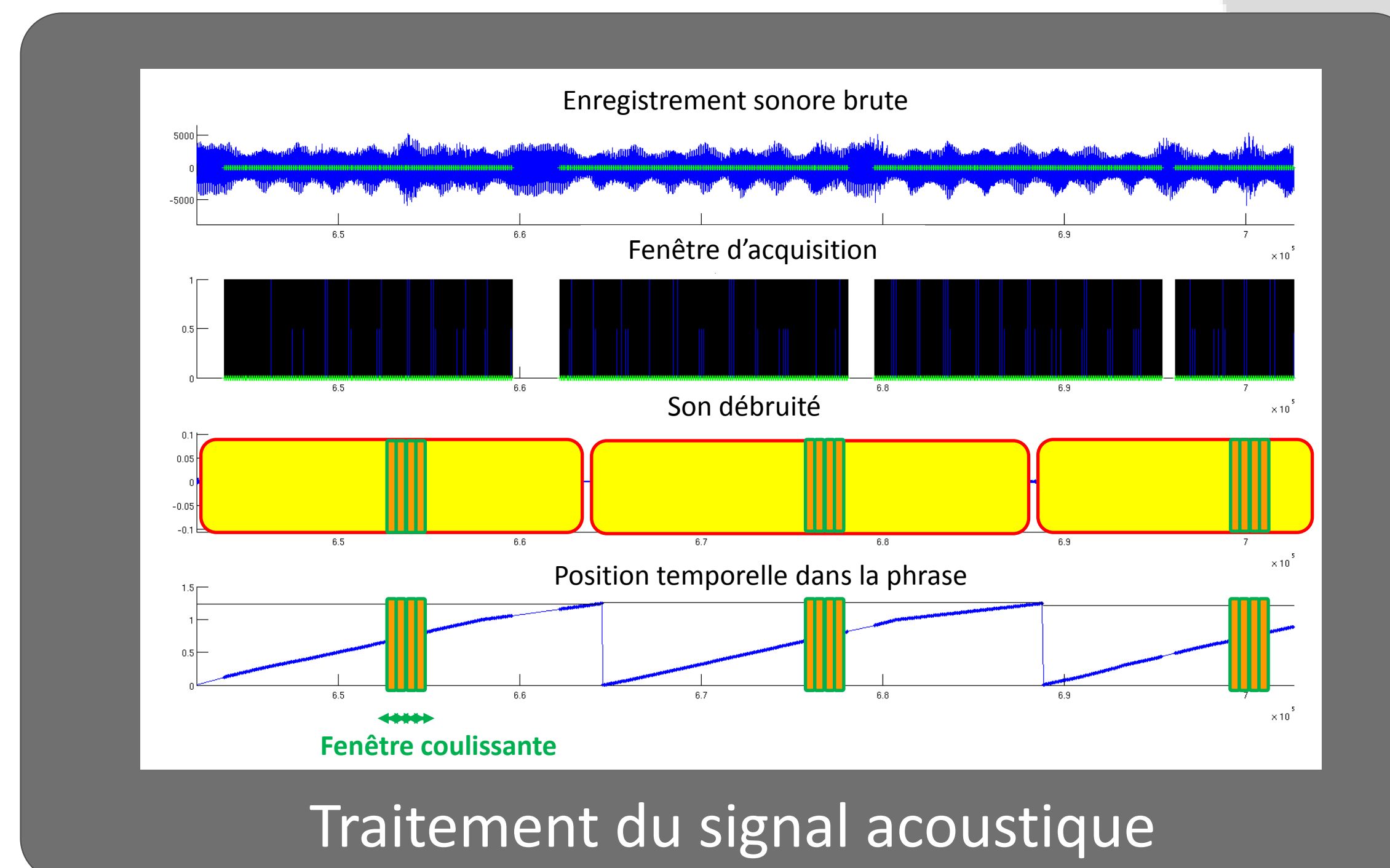
- Séquence b-SSFP, 1 coupe sagittale, matrice 256x256, TR/TE = 3,9/1,7ms, épaisseur 5mm, angle de bascule 45°, Champ de vue 30cm, 65 phases temporelles, durée d'acquisition environ 1 min.
- Dans le protocole, dix courtes phrases ont été sélectionnées afin de couvrir la variabilité articulatoire française [3] : « Ma chemise est roussie », « Donne un petit coup », « Voilà des bougies », « Une réponse ambiguë », « Louis pense à ça », « Mets tes beaux habits », « Une pâte à choux », « Prête-lui seize écus », « Chevalier du gué », « Il fume son tabac ». Le sujet a dû répéter une des dix phrases pendant toute l'acquisition:
- Le son a été enregistré en utilisant un microphone optique (FOMRI III, Optoacoustics, Yehuda Israel). Le signal d'ouverture des fenêtres d'acquisition a aussi été enregistré pour permettre la synchronisation des événements acoustique et IRM.

➤ Traitement du signal acoustique

- L'enregistrement sonore a été débruité [4] pour éliminé le bruit de la commutation du fourreau de gradient de champ magnétique;
- Une segmentation manuelle de chaque phonème lors du dépouillement de l'enregistrement;
- Un signal de phase acoustique a été généré pour indiquer la position temporelle de l'acquisition de chaque ligne de l'espace-k à l'intérieur de la phrase.

➤ Reconstruction de la boucle Cinématique

- Une normalisation linéaire par parties basée que la segmentation manuelle a été utilisée pour prendre en compte la variabilité des répétitions des occurrences de chaque phrases;
- Les boucles cinématiques pour chacune des phrases ont été reconstruites en utilisant l'algorithme compensé en mouvement cine-GRICS [5] : une fenêtre coulissante de 80 ms a été choisie et la phrase a été reconstruite en 128 images (résolution temporelle <10ms).



Résultats

- En Figure 1 deux images du sujet prononçant « Voilà des bougies » qui montre l'absence de flou ou d'artéfact de mouvement dans la boucle cinématique reconstruite.
- En Figure 2, la variation de distance d'une part entre le dos de la langue et la voûte palatine et d'autre part entre le dos de la langue et la paroi pharyngale est mise en corrélation avec l'enregistrement sonore de la phrase.

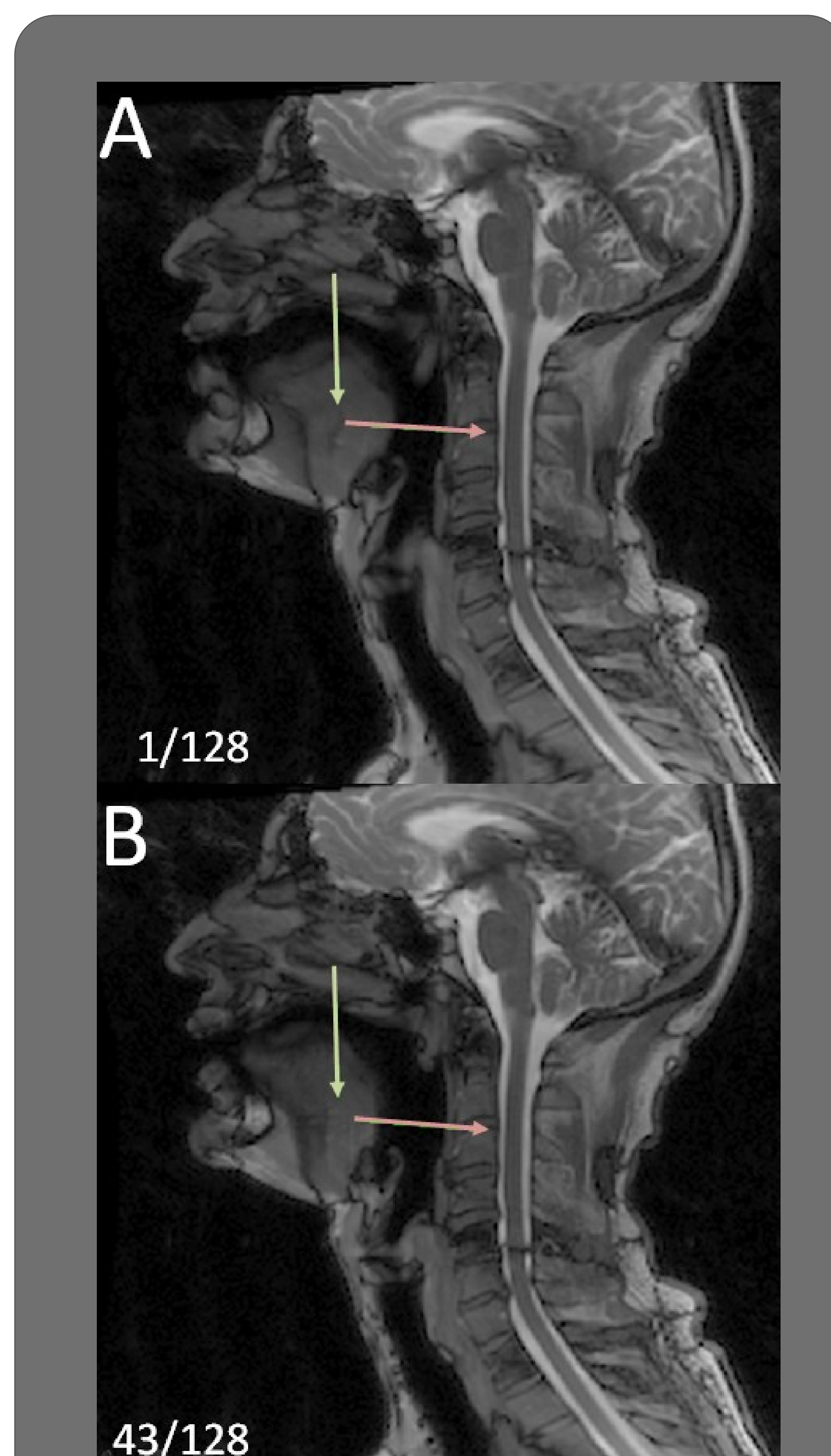


Fig 1. : Deux positions temporelles (A) 1, (B) 43 parmi les 128 images de la boucle ciné en coupe sagittale le long du conduit vocal avec les positions des modes TM annotées.

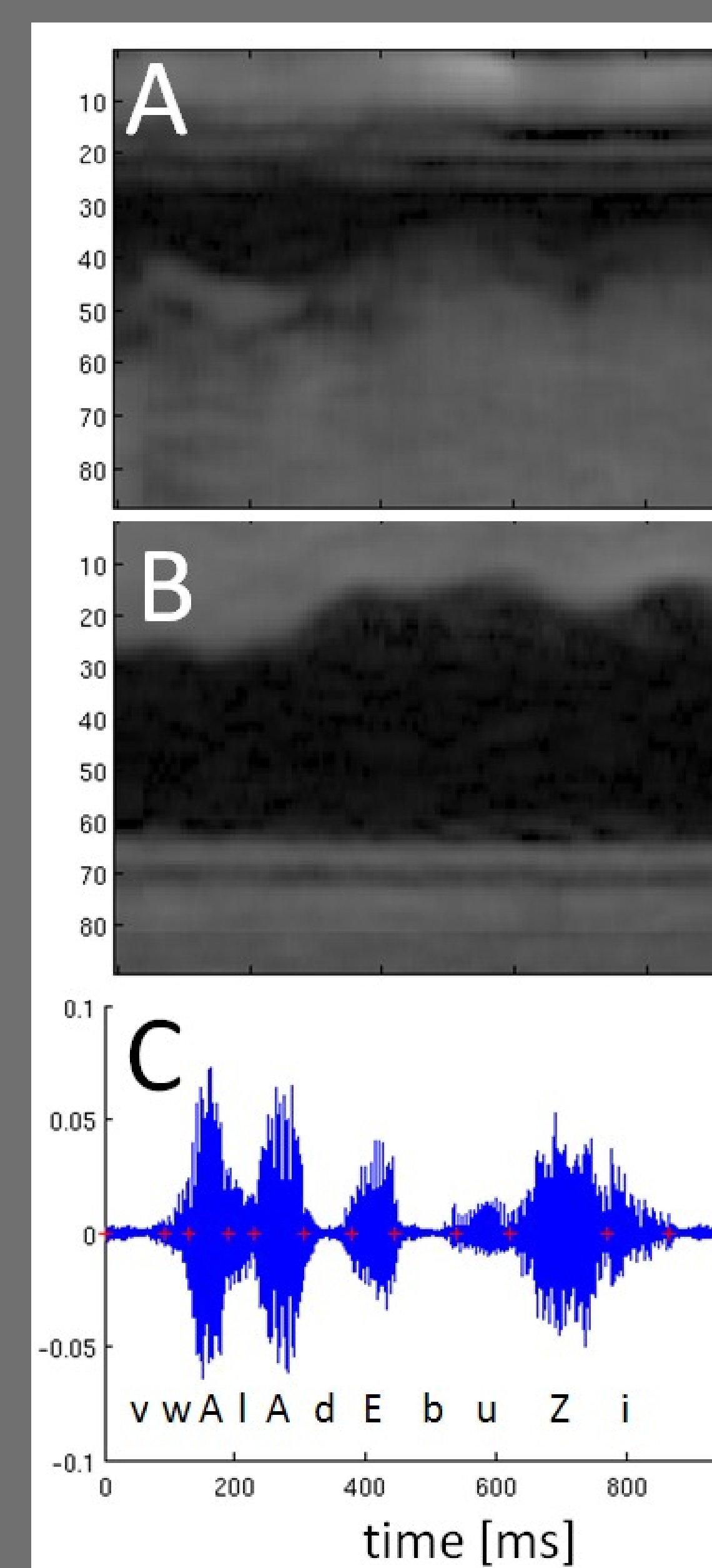


Fig 2. : Visualisation temps mouvement de la boucle ciné pour (A) de la voûte palatine au dos de la langue et (B) du dos de la langue à la paroi postérieur pharyngale, (C) enregistrement sonore utilisé comme signal de synchronisation dans l'algorithme de reconstruction afin de compenser le mouvement.

Discussion et conclusion

- Un problème potentiel de l'utilisation de la séquence SSFP est la présence d'artéfacts de bande causés par les variations intenses de B0 aux interfaces air-tissu.
- Chaque boucle cinématique permet une segmentation du conduit vocal avec une résolution spatiale et temporelle suffisante. Ainsi, un examen IRM d'une demi-heure permet l'acquisition d'un modèle vocal personnalisé.

Références

- [1] Narayanan et al., J Acoust Soc Am, 115(4):1771 (2004) ; [2] Frauenrath et al., Act Acus, 94(1):148 (2008);[3] Maeda, Actes X JEP, p152, Grenoble (1979) ; [4] Ozerov et al., IEEE TASLP, 20(4) :1118 (2012); [5] Vuissoz et al., JMIR, 35 :340 (2012) .