

# **RAPPORT PROJET**

## **Modélisation et Calcul Scientifique**

BRODERS NICOLAS – CAUSSÉ EMMANUEL  
COURDY-BAHSOUN CLÉMENCE – COUSIN VINCENT

M1 IHM – M1 SIAME – M1 IGAI – M1 DC

23.11.2018

# Rapport

# 1 Application d'enregistrement de commande vocale

## Justification d'implémentation :

Le choix d'extension de classe entre *AudioRecord* et *MediaRecorder* pour l'enregistrement se justifie par l'exigence que l'audio enregistré soit au format **WAV**. En effet, même si *MediaRecorder* permet d'enregistrer des fichiers audio sous plusieurs formats assez rapidement, il n'intègre pas dans ces méthodes l'enregistrement au format WAV. Il aura donc fallu partir de la classe *AudioRecord* et de faire quelque modification (ajout d'un EN-TÊTE WAV) pour obtenir le dit format.

## Optimisation possible :

Même si la seule exigence d'enregistrer au format WAV a été remplie, il existe des améliorations pour rendre plus agréable la navigation au sein même de l'application, à savoir :

Côté technique : – A la fin d'un enregistrement, armer le lecteur audio sur le dernier audio enregistré. – Si on accède à la liste des audios enregistrés, armer le lecteur audio sur le premier enregistrement.

Côté graphique : – La possibilité de visualiser la force de la voix lors de l'enregistrement avec une courbe permettrait à l'utilisateur d'avoir un feedback en direct et lui permettrait de s'ajuster. (Les smartphones n'ont pas la même sensibilité au niveau du micro) Pour plus d'information le Manuel Utilisateur est disponible en Annexe.

## 2 DTW sans ACP

Cette partie correspond aux résultats obtenus en terme de taux de reconnaissance par l'implémentation de la méthode DTW dans la première partie du TP. Dans laquelle on a pu constater que l'enregistrement de la base d'apprentissage avait un effet favorable sur la reconnaissance vocale.

### Résultats :

Matrice de confusion sur un jeu de test de voix féminines avec une base d'apprentissage enrichie par des ordres prononcés par un accent et avec application du DTW classique

$$\begin{bmatrix} 5.0 & 0.0 & 0.0 & 0.0 & 5.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 6.0 & 0.0 & 0.0 & 0.0 & 0.0 & 6.0 & 0.0 \\ 0.0 & 0.0 & 6.0 & 0.0 & 0.0 & 6.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 3.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 2.0 \end{bmatrix}$$

Le taux d'erreur calculé est d'environ 65,625%

### 3 DTW avec ACP

L'**Analyse en Composants Principals** est une méthode permettant ici de représenter les enregistrements par des points dans un repère orthonormé en 3 dimensions. Ainsi on réduit la taille des données à traiter en moyennant les Fields de chaque enregistrement, puis en les représentant sous forme d'un point dans le repère défini grâce aux vecteurs propres des plus grandes valeurs propre de la matrice de variance-covariance obtenu grâce à la base d'apprentissage.

**Résultats** Matrice de confusion pour les mêmes données, en appliquant l'ACP, obtenues en comparant la distance entre chacun le point de l'enregistrement avec chacun des points de la base d'apprentissage.

$$\begin{bmatrix} 2.0 & 0.0 & 0.0 & 0.0 & 2.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 6.0 & 4.0 & 3.0 & 0.0 & 4.0 & 6.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 2.0 & 0.0 & 0.0 & 2.0 & 0.0 & 0.0 \\ 3.0 & 0.0 & 0.0 & 0.0 & 3.0 & 0.0 & 0.0 & 2.0 \end{bmatrix}$$

Le taux d'erreur calculé est d'environ 89,0625%

## 4 Méthode des k-ppv

La méthode des k-plus proches voisins consiste à chercher parmi l'ensemble des points obtenus par croisement des résultats de l'ACP avec les enregistrements moyennés de la base d'apprentissage, les k enregistrements les plus proches en distance. Pour cela on a arbitrairement fixé k au nombre d'ordres différents enregistrés dans la base d'apprentissage. Nous avons également utilisé une map afin de labeliser chaque point par l'ordre qu'il doit reconnaître. Ainsi les k points les plus proches sont classés selon l'ordre qu'il reconnaisse, et c'est la classe majoritaire qui donnera l'ordre reconnu.

### Résultats :

Matrice de confusion sur les mêmes données obtenues par la méthode des k-ppv basée sur l'ACP

$$\begin{bmatrix} 4.0 & 0.0 & 0.0 & 0.0 & 4.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 2.0 & 3.0 & 2.0 & 0.0 & 3.0 & 2.0 & 0.0 \\ 1.0 & 2.0 & 2.0 & 0.0 & 1.0 & 2.0 & 2.0 & 2.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 2.0 & 1.0 & 1.0 & 0.0 & 1.0 & 2.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}$$

Le taux d'erreur calculé est d'environ 84,375%

## 5 Conclusion

### Resultat

Les résultats des différentes analyses nous montre que la reconnaissance basée sur l'ACP a un taux d'erreur plus élevé que la reconnaissance par DTW. Ces résultats semble cohérent, étant donné que l'ACP vise à réduire le nombre et la taille des données. Contrairement à DTW qui va donner la distance en parcourant toutes les fenêtres des deux enregistrements, l'ACP permet de calculer la distance entre deux points dans un repère. Cependant l'obtention de ces points nécessite de moyenner les données de chaque field pour obtenir un vecteur dans  $\mathcal{R}^{13}$ . Ainsi le taux d'erreur peut être plus élevé, néanmoins les calculs sont sensés être moins coûteux.

On remarque que sur le test mené, la reconnaissance par les k-ppv est meilleure. En effet on augmente le taux de reconnaissance car on va chercher parmi tous les voisins proches la classe la plus représentative. Pour obtenir des résultats plus significatif, il aurait fallu enrichir la base d'apprentissage de manière à obtenir un nuage de point beaucoup plus fourni. En théorie les enregistrements d'un même ordre devrait recouvrir une zone du repère, alors plus il y a d'enregistrement d'un ordre dans la base de référence et plus cette zone se précisera, ainsi le taux d'erreur devrait diminuer encore.

### Remarque

L'affichage du nuage de point de la base d'apprentissage aurait été un plus, cependant après quelques recherches sur la manière d'obtenir un repère en 3 dimensions avec l'ensemble des points à partir de java, il s'est avéré qu'il fallait utiliser la bibliothèque java3d, ou une bibliothèque équivalente, basée sur opengl. Le manque de maîtrise de ce genre de bibliothèque actuellement, et un certain manque de temps explique notre choix de ne pas fournir l'affichage des points. Il existait tout de même une alternative qui nous aurait pris un certain temps en créant nous même le repère et en plaçant les points manuellement. Néanmoins le graphe n'aurait pas pu être recalculé à chaque fois que la base d'apprentissage aurait été enrichie.

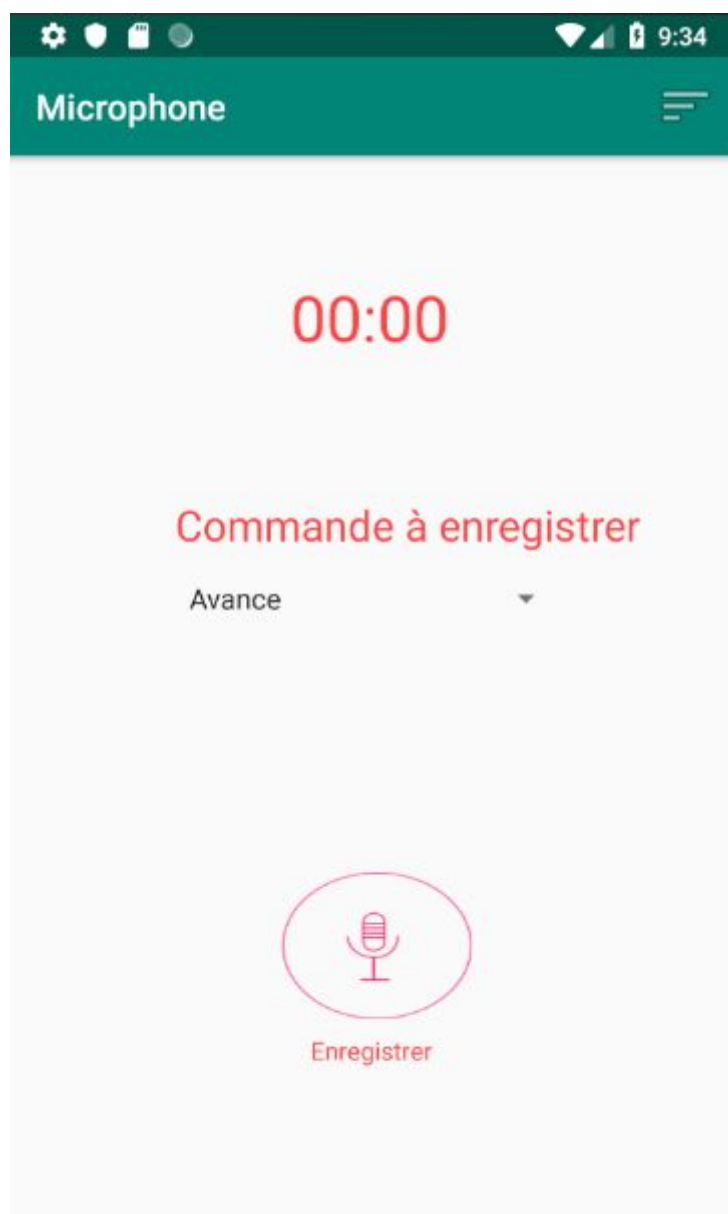
# Annexe



# Application Microphone

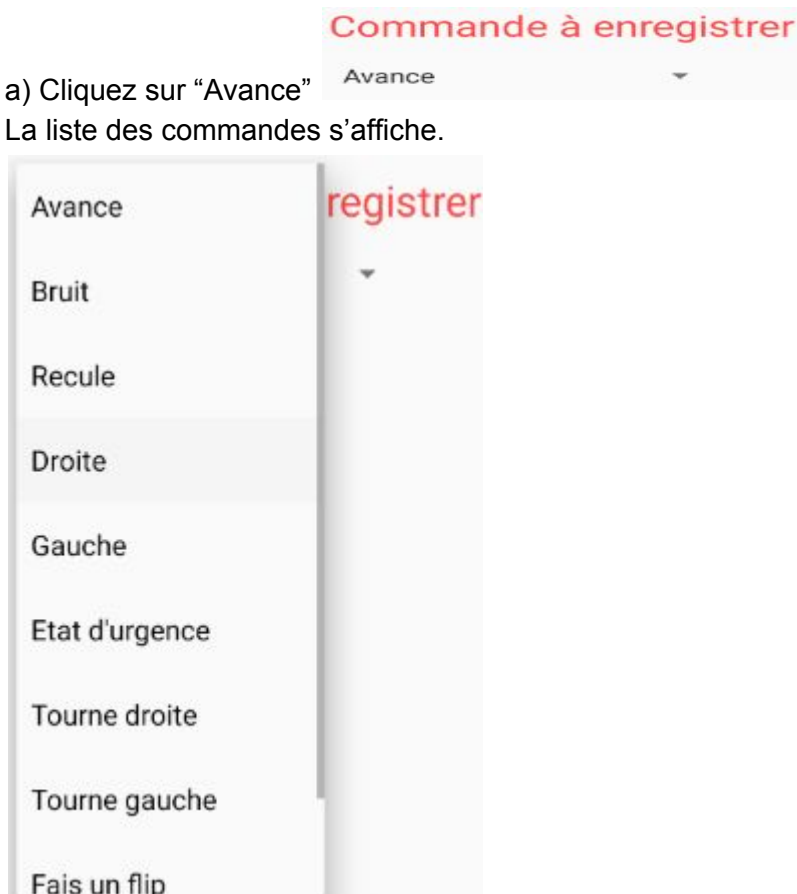
## Manuel utilisateur

### I - Enregistrer une commande audio



- 1) a) Cliquez sur “Avance”

La liste des commandes s’affiche.

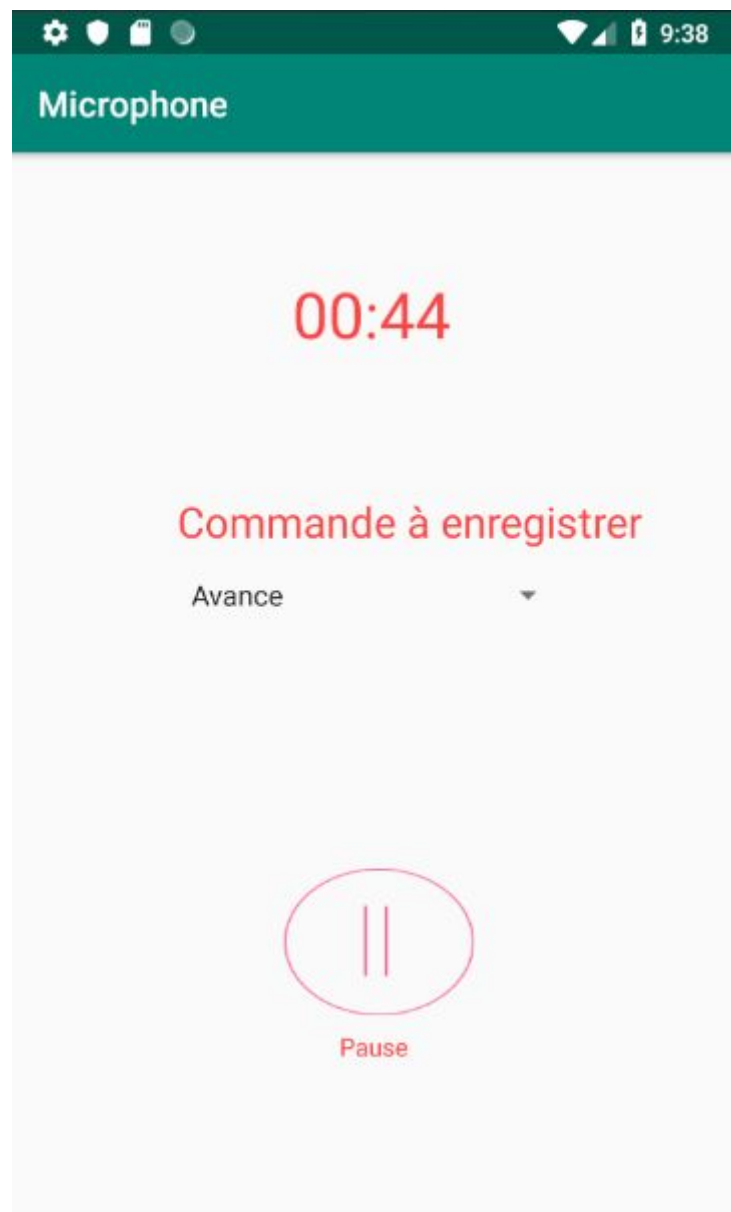


- b) Sélectionnez la commande à enregistrer dans la liste.

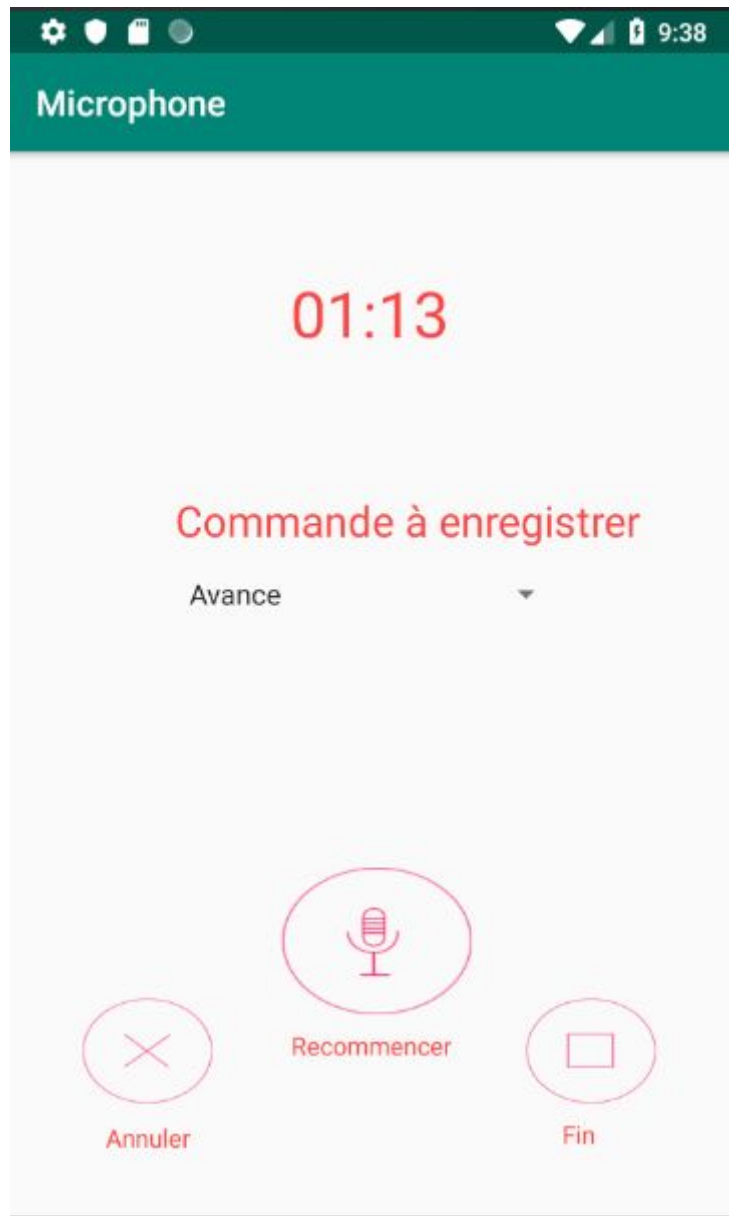
- 2) Appuyez sur le bouton



pour commencer à enregistrer.




- 3) Appuyez sur le bouton  pour arrêter l'enregistrement.




4) Appuyez sur le bouton  pour terminer d'enregistrer.

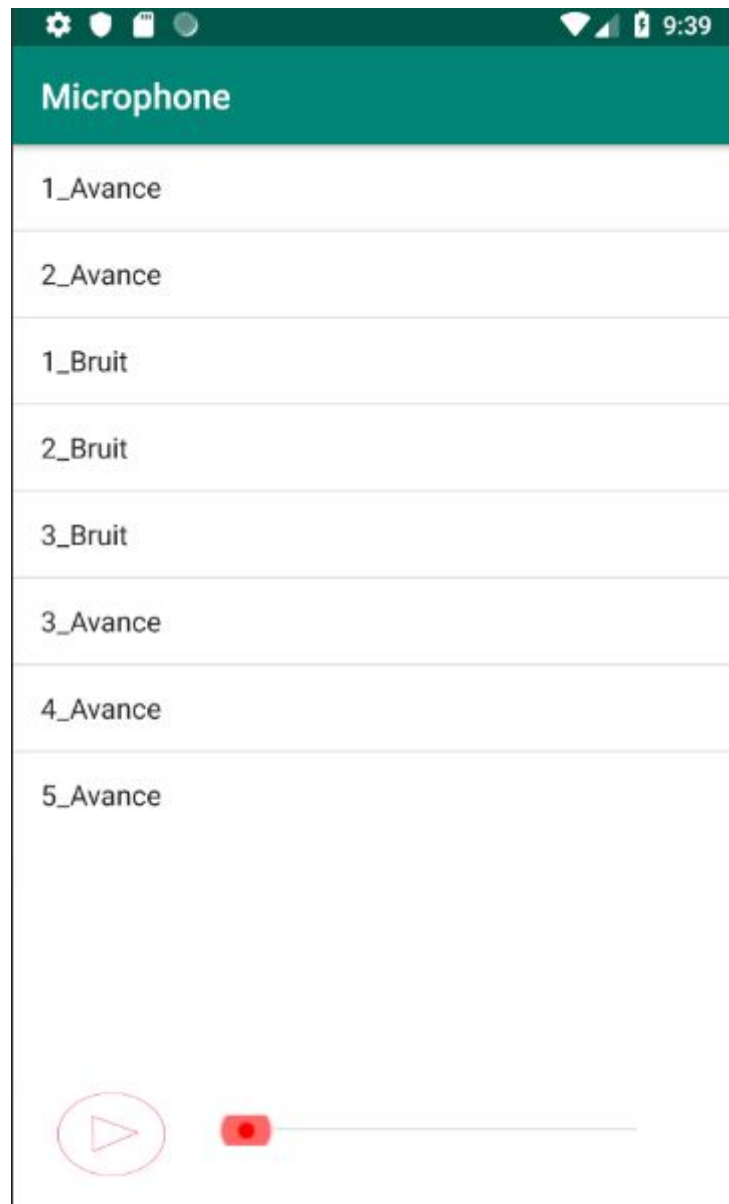
Vous pouvez annuler l'enregistrement en appuyant sur le bouton 

Vous pouvez recommencer directement l'enregistrement en

appuyant sur le bouton 

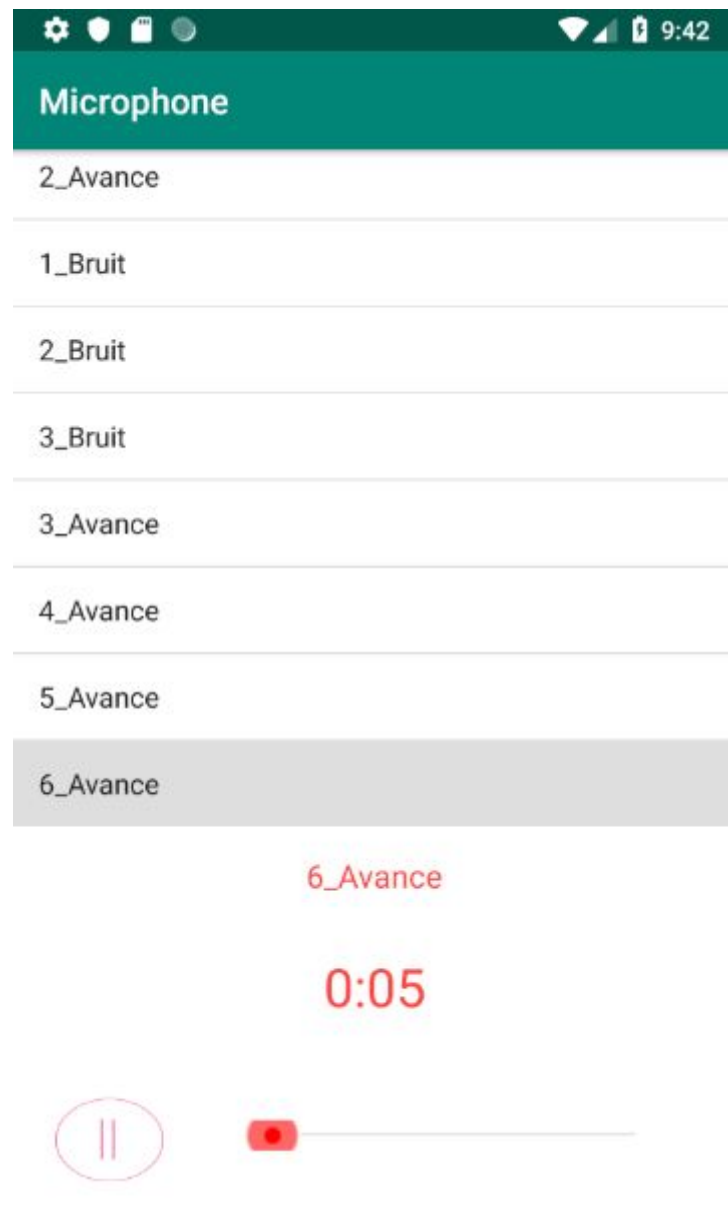
## II - Accéder à la liste des audio enregistrés

- 1) Cliquez sur le bouton  sur la page d'accueil.  
Si celui-ci ne s'affiche pas, veuillez à ce qu'aucun enregistrement ne soit en cours.



### III - Lire une commande audio

- 1) Cliquez sur la commande correspondante dans la liste.



Vous pouvez faire progresser la lecture de l'enregistrement en déplaçant le curseur vers la gauche ou vers la droite



### III - Supprimer une commande audio

- 1) Appuyez longuement sur l'audio que vous souhaitez supprimer
- 2) Cliquez sur supprimer.

ATTENTION : Ceci aura pour effet de supprimer l'audio du téléphone de manière définitive.

