

MANUEL D'UTILISATION DU COMPAREUR

Version 2 (PC)

Abderrahim Meqqori, Jacques Durand et Coralie Meynier

Introduction

Le but de ce manuel est de décrire la Version 2 du Compareur mis au point par Abderrahim Meqqori pour le projet PFC (Phonologie du Français Contemporain). Cet outil, comme son nom l'indique, permet de comparer des réalisations sonores (comme les mots de la liste du projet PFC) par différents locuteurs. Dans sa version actuelle, le compareur s'applique à toute liste de mots disponible sous deux formes : un fichier .wav et un fichier Textgrid issu d'une segmentation dans Praat. Cette segmentation, pour que l'outil fonctionne parfaitement, doit obéir aux règles établies dans PFC. Il faut que le fichier .wav ait été épuré : exclusion des remarques précédant ou suivant l'enregistrement de la liste, exclusion de remarques parenthétiques, bornes positionnées correctement sans intervalles vides, pas de borne initiale ni de borne finale (« 1 roc » commence à zéro, et « 94 brin » finit le fichier). La transcription alignée sous Praat est du type : 1 roc, 2 rat, etc. qui correspond à des fichiers Textgrid comme le suivant :

```
File type = "ooTextFile"
Object class = "TextGrid"
xmin = 0
xmax = 131.34462585034012
tiers? <exists>
size = 1
item []:
  item [1]:
    class = "IntervalTier"
    name = "g.textgrid"
    xmin = 0
    xmax = 131.34462585034012
    intervals: size = 94
    intervals [1]:
      xmin = 0
      xmax = 1.381619455889517
      text = "1 roc"
    intervals [2]:
      xmin = 1.381619455889517
      xmax = 2.683009716617327
      text = "2 rat"
```

1. Préparation des données

Avant d'utiliser le Compareur, il faut s'assurer que le logiciel Quicktime (Apple) est bien installé sur la machine. Il est libre et récupérable sur Apple.com.

Pour utiliser le Compareur, il faut créer un dossier dans lequel on placera l'application elle-même, un fichier texte *listemots.txt* contenant la liste de mots sous forme brute (a) et un fichier *base.txt* (b).

a. Exemple d'un début de fichier texte *listemots.txt* :

```
1   roc
2   rat
3   jeune
4   mal
```

5 ras

Ce fichier texte entier correspondant à la liste officielle PFC est fourni par commodité avec le Comparateur. Il ne doit pas comporter de titre ou d'annotations. Il doit être enregistré au format texte avec l'extension *.txt*. On peut utiliser toute autre liste de mots à condition de respecter le format. Cela permet de traiter toute liste complémentaire conçue au sein du projet.

Le deuxième composant obligatoire de cet outil est la construction d'une base d'information qui, on le verra, peut rester extrêmement simple. Un exemple est donné ci-dessous

b. Exemple de fichier texte *base.txt* :

```
11AMG2MW@Femme@48@1953@Femme de ménage@Bac @Douzens@5/5@Occitan:moyen
11ANB1MW@Femme@48@1953@Femme de ménage@Bac @Douzens@5/5@Occitan:moyen
11AML1MW@Femme@75@1926@Retraitée@Brevet@Douzens@5/5@Occitan:courant
11AMG1MW@Femme@18@1983@Etudiante@Bac @Douzens@4/5@Occitan:non
```

L'élément incontournable de *base.txt* est la première information de chaque paragraphe. En effet, c'est à partir des noms de locuteurs + MW que le Comparateur retrouve les fichiers *.wav* et *Textgrid* à traiter. Ces noms doivent obligatoirement répondre aux normes PFC, à savoir 8 lettres dont les deux dernières sont MW. Une base minimale utilisable sur le comparateur pourrait ne contenir que les noms des fichiers *.wav* :

```
11AMG2MW
11ANB1MW
11AML1MW
11AMG1MW
```

Dans la mesure où le Comparateur fournit des informations « biographiques » sur les locuteurs, la base doit être construite selon le modèle suivant si l'on tient à afficher ces informations et à utiliser la recherche croisée :

Modèle pour *base.txt* :

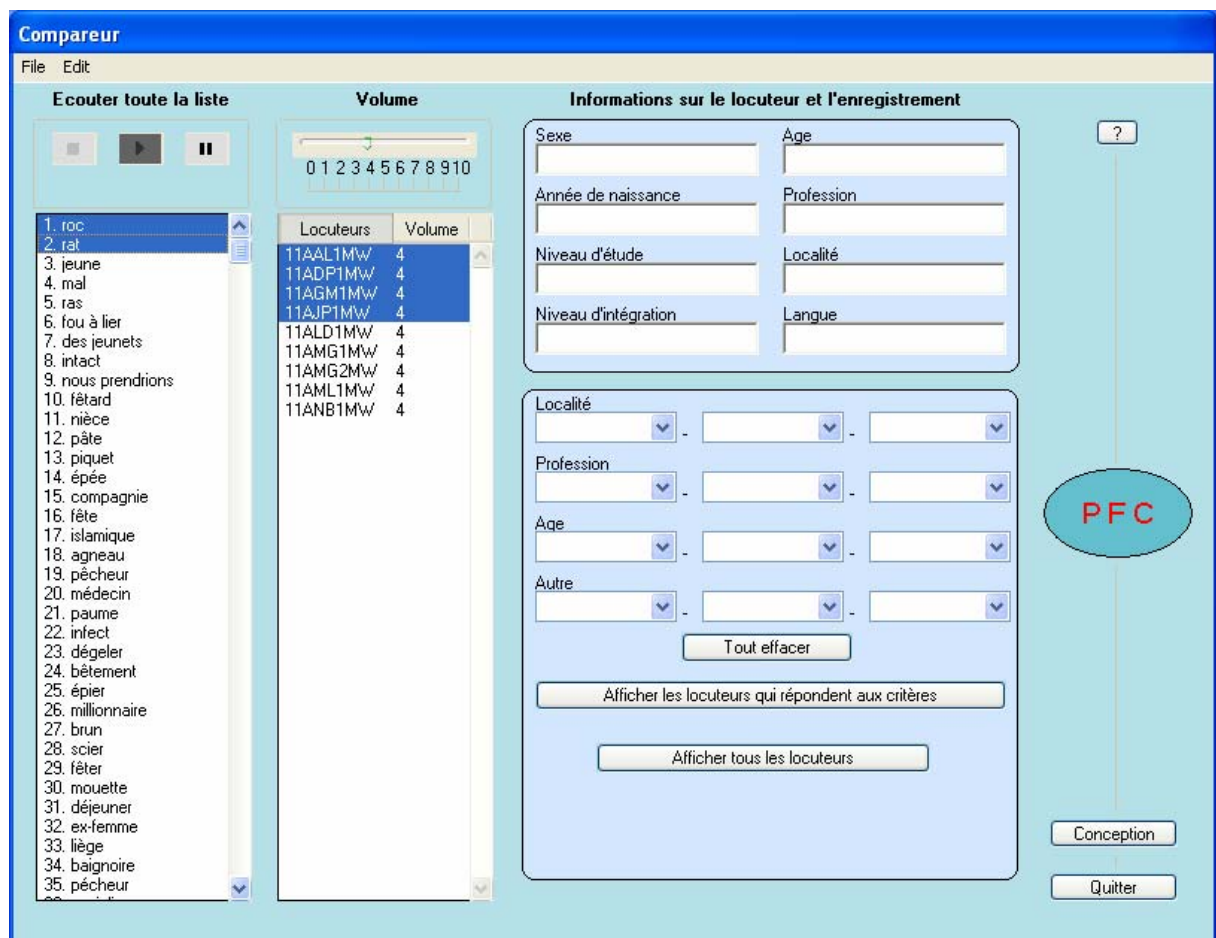
```
Nom_du_fichier_wave@Sexe@Age_au_moment_de_l'enquête@Année_de_naissance@Pr
ofession@Niveau_d'études@Point_d'enquête@Degré_d'intégration@Autre_langue
_pertinente
```

2. Utilisation du comparateur

Le Comparateur est un fichier exécutable. A son démarrage s'affiche la fenêtre standard d'ouverture de document sous Windows :



A ce stade, il faut chercher le dossier du point d'enquête qu'on veut traiter. Une fois ce dernier localisé, il convient de le sélectionner en cliquant dessus et de cliquer sur le bouton **OK**. On accède alors à la fenêtre principale du Compareur :



Le premier champ reproduit la liste des mots de la liste qui a été saisie dans le fichier *listemots* constitué au départ. Le deuxième champ fournit la liste des locuteurs contenus dans le dossier du point d'enquête choisi. Surplombant ce deuxième champ, on peut observer une barre de réglage qui permet d'ajuster le volume sonore pour chaque locuteur individuellement. La valeur de ce réglage qui est à 4 par défaut apparaît en face de chaque locuteur. Un ajustement individualisé est habituellement nécessaire dans la mesure où tous les enregistrements sont disparates du point de vue acoustique.

La fonction du Comparateur est double. Il permet en premier lieu une écoute et une comparaison de la prononciation de mots réalisés par différents locuteurs. Une fois un mot sélectionné, un simple clic sur tel ou tel locuteur permet de l'écouter. Cette version du Comparateur permet une sélection multiple aussi bien des mots que des locuteurs. Autrement dit, on peut comparer, avec une seule sélection, plusieurs locuteurs pour un mot ou plusieurs mots pour un seul locuteur ; il suffit pour cela de sélectionner le nombre souhaité de mots et de locuteurs. En revanche, la sélection de plusieurs mots pour plusieurs locuteurs est instable. Il est déconseillé d'y recourir.

Cet outil ne peut pas se substituer à une étude acoustique mais, d'un point de vue auditif, offre un moyen de comparaison rapide et confortable. En deuxième lieu, il permet, pour un locuteur donné, d'écouter individuellement et aléatoirement les mots de la liste. Une fois un locuteur sélectionné, un simple clic sur tel ou tel mot permet de l'écouter.

Pour avoir des conditions d'écoute optimale, notre expérience suggère d'identifier le locuteur pour lequel le volume est le plus haut, d'ajuster le volume des haut-parleurs externes sur ce dernier et ensuite de monter individuellement le volume des autres locuteurs.

Un ensemble de champs situés en haut à droite de la fenêtre permet de visualiser les informations qui ont été saisies dans le fichier *base.txt*. (Localité, âge, sexe, etc.)

Sous ce cadre, un ensemble de menus déroulant permet d'effectuer des recherches multicritères. On a ainsi la possibilité de n'afficher, pour comparaison, que les locuteurs répondant aux critères sélectionnés. On peut par exemple opter pour une comparaison des locuteurs de Toulouse et de Bordeaux, ayant entre 25 et 35 ans, etc.

Bon travail !