

# Corpus OFROM

## État du corpus au 1er avril 2015

La base de données a été mise à jour et compte, au premier avril 2015, un total de 448'798 mots (42 heures) répartis sur 197 locuteurs.

Le détail des changements, avec la liste des transcriptions ajoutées, la liste des corrections apportées au corpus et toute remarque utile, est présenté ci-dessous.

### Ajouts (transcriptions) :

unine15-001	unine15-002	unine15-003	unine15-006
unine15-007	unine15-008	unine15-009	unine15-010
unine15-011	unine15-012	unine15-014	unine15-015
unine15-016	unine15-017	unine15-018	unine15-020
unine15-021	unine15-022	unine15-023	unine15-024
unine15-025	unine15-026	unine15-027	unine15-028
unine15-029	unine15-030	unine15-031	unine15-033
unine15-034	unine15-035	unine15-036	unine15-037
unine15-901	unine15-902	unine15-903	unine15-904
unine15-905	unine15-906		

### Corrections :

- Les « ne » de négation qui ne sont pas prononcés ont été supprimés. Dans le cas ambigu de « on n'a pas », le « n' » est conservé.
- Le « il » prononcé /i/ a été systématiquement écrit « il ». Exemple : /kãtidi/ est écrit « quand il dit ». Dans le cas de « il y a » prononcé /ja/, on considère que le « il » n'est pas prononcé.
- Le symbole pour l'amorce a été changé en « / ». Exemple : « qui habi/ qui qui ».

### Remarques :

Le document a été réalisé après-coup, courant mai 2015.

Avant cette date, aucun journal des corrections apportées au corpus n'avait été établi. D'autres corrections de détail ont pu survenir, qui concernent la transcription ou les métadonnées.

Les ajouts n'indiquent que les locuteurs : un locuteur pouvant avoir plus d'une transcription, le rapport manque de précision.

(août 2016) Le calcul du nombre de mots / durée a changé. L'ancien rapport indiquait 388'488 mots.