

Corpus OFROM

État du corpus au 1er août 2015

La base de données a été mise à jour et compte, au premier août 2015, un total de 583'240 mots (54 heures) répartis sur 222 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) :

unine15-013	unine15-019	unine15-038	unine15-039
unine15-040	unine15-041	unine15-042	unine15-043
unine15-045	unine15-046	unine15-047	unine15-048
unine15-049	unine15-052	unine15-053	unine15-054
unine15-055	unine15-056	unine15-057	unine15-060
unine15-061	unine15-062	unine15-063	unine15-064
unine15-065			
	unine15-915	unine15-916	unine15-917
unine15-918	unine15-924	unine15-925	unine15-927
unine15-928			

Corrections :

- 4 juin 2015 : généralisation de « oh là là » dans unine11b01m, unine11b17m et unine11c05m.
- Correction des chevauchements
- Corrections de détail dans unifr11a02m, unifr14a05d, unine08a16m, unine08a26m, unine08a30m, unine11a03d, unine11b14m, unine11b20m, unine12a07m, unine12a10m, unine13a03d, unine15a04d, unine15a37d, unine15a40d
- 2 juillet 2015 : généralisation de « tu » élidé devant voyelle (quand réalisé) uniquement.
- Suppression des transcriptions pour unine11-rsb et unine12-asa.
- 09 juillet 2015 : généralisation de « / », pour les unine12 notamment.
- 24 juillet 2015 : corrigé les chevauchements dans unine15a08d

Remarques :

Le suivi des corrections n'a pas été fait : les dates peuvent être imprécises et certaines corrections de détail peuvent manquer.

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 615'621 mots et 64 heures.

François Delafontaine
francois.delafontaine@unine.ch