



OFROM

Conventions de transcription

*Ce document n'est pas exhaustif.
Pour de plus amples et complètes informations, consulter le document de référence (p.6-10).*

Principes adoptés dans le corpus

- 1) Emploi de l'orthographe standard.
- 2) Transcription uniquement de ce qui est entendu.
ex. : | pense que ils mangent pas tellement bien mais bon | (unine13a01d, unine13-hma¹)
- 3) Pas de ponctuation.

Signification des symboles employés dans le corpus

- | : frontière d'un segment sonore transcrit (unité de base pour la recherche)
- _ : pause (peut contenir rires, bruits, etc.)
- % : segment inaudible
- @ : propos de l'enquêteur ou de locuteurs tiers
- # : parties anonymisées
- / : troncation

Quelques conventions utiles à la recherche

- Ne** : seulement transcrit si la négation est audible ou après le phonème /n/
ex. : | autrement on n'avancait pas | (unine15a40d, unine15-028)
- Puis** : seulement transcrit lorsque le /i/ est clairement audible, sinon transcrit « pis »
- Il, ils** : toujours transcrit « il », « ils » (même si on entend /i/)
- Il y a** : toujours transcrit « il y a »
N.B. : avant janvier 2020, transcrit « y a » si le « il » était inaudible
- Tu as, tu es** : toujours transcrit « tu as », « tu es »
N.B. : avant janvier 2020, transcrit « t'as », « t'es » lorsque le /y/ était inaudible)
- Nombres** : écrits en toutes lettres ; emploi des règles standards : tiret en-dessous de la centaine (ex. : « trente-cinq mille trois cents »)
- Onomatopées** : exemples fréquents : « ah », « ben », « eh », « euh », « euh mh », « mh », « oh ».
N.B. : certains cas plus rares sont considérés comme des segments inaudibles et sont transcrits « % » (ex. : « pfou »).

¹ Identifiant de l'enregistrement, identifiant du témoin