



UMR 5191 - CNRS / Université Lyon 2
Interactions, Corpus, Apprentissages, Représentations

Convention ICOR

Mise à jour : janvier 2013

Groupe ICOR

UMR 5191 ICAR (CNRS – Lyon 2 – ENS de Lyon)



Avis au transcripteur :

Le présent module de la convention ICOR sert à la notation des phénomènes verbaux et vocaux. Les phénomènes non verbaux (action non verbales, gestes, mimiques, etc.) font partie d'un autre module de convention.

Chaque corpus et chaque projet de recherche ont leurs propres exigences de transcription. Les usages de la Convention ICOR ne seront donc pas toujours les mêmes. Il est recommandé (et pour l'intégration d'un corpus dans la base CLAPI, il est nécessaire) de spécifier ces usages selon les trois niveaux suivants :

- *niveau de base* : y sont obligatoirement notés les changements de locuteur + les chevauchements + les pauses (indiquer la manière dont les pauses ont été mesurées : soit par utilisation d'un logiciel, soit par estimation perceptuelle) ;
- *niveau intermédiaire* : sont ajoutés au niveau de base les notations des phénomènes vocaux et les commentaires ;
- *niveau spécialisé* : les spécificités apportées à ce niveau de transcription sont définies par la thématique de recherche suivie, par exemple étude sur les pratiques de liaison, ou sur l'emploi de mots étrangers, etc.

Il convient également de fournir une liste des phénomènes supplémentaires notés dans votre transcription et non répertoriés dans la présente convention.

Lors de l'intégration d'un corpus dans la base CLAPI, ces spécifications sont à joindre dans un document séparé de celui de la transcription et des autres éléments du corpus.

Attention :

Avant toute chose, **désactivez vos mises en forme / remplacements automatiques** (espaces insécables, insert d'espaces avant les deux points) ainsi que les majuscules.

L'**alignement gauche** des paragraphes est le seul autorisé, l'alignement justifié est par conséquent interdit.

La police à utiliser pour les transcriptions est **Courier taille 10**, parce que tous les caractères y ont la même largeur.

Toutes les productions verbales sont transcrites en **minuscule** ; les majuscules sont réservées pour la notation des saillances perceptuelles.

L'usage de tabulations est exclu sauf après la notation de l'identité du participant. Pour aligner les chevauchements, utiliser des espaces car ils sont les seuls à être conservés lors des transferts de fichier entre ordinateurs.

Anonymisation :

Remplacer tous les noms propres (personnes, lieux...) par des pseudonymes, en tentant de respecter le nombre de syllabes et les éventuelles connotations (prénom américain...).

Remplacer aussi les chiffres, nombres et numéros (n° de tel, de rue, de compte bancaire...).

Pour des informations plus détaillées sur l'anonymisation, consulter le site CORINTE : <http://icar.univ-lyon2.fr/projets/corinte/>, menu "Questions juridiques".

Table des matières

1 Identité du participant

- 1a Participant identifié
- 1b Participant incertain
- 1c Hésitation entre deux participants
- 1d Participant non identifié
- 1e Événement non attribuable à un participant

2 Tour de parole

- 2a Notation du tour
- 2b Enchaînement immédiat (latching)
- 2c Chevauchement

3 Actions

- 3a Action à valeur de tour
- 3b Action non attribuable à un participant

4 Structures segmentales (transcrire le verbal)

- 4a Orthographe
- 4b Elision standard et non standard
- 4c Allongement
- 4d Amorce de mot
- 4e Liaison facultative faite
- 4f Liaison obligatoire entre deux mots non faite
- 4g Nombres
- 4h Acronymes, sigles

- 4i Noms propres
- 4j Onomatopées
- 4k Segments inaudibles dont le nombre de syllabes est identifiable
- 4l Segments inaudibles dont le nombre de syllabes n'est pas identifiable
- 4m Multitranscriptions d'un même segment et Segments incertains

5 Pauses intra-tour

- 5a Micro pause
- 5b Pause chronométrée

6 Productions vocales

- 6b Aspiration
- 6c Expiration, soupir

7 Prosodie

- 7a Montée et chute intonatives
- 7b Saillance perceptuelle

8 Descriptions

- 8a Généralités
- 8b Passages non transcrits
- 8c Alternance de codes

9 Liste des onomatopées, interjections et autres particules

	Phénomène	Conventions en format liste	Exemples en format liste
1- Identité du participant (tire de transcription par locuteur)			
1a	Participant identifié	Identifiant en début de paragraphe du tour (voir 2a). L'identifiant est composé d'un, de deux ou de trois caractères (en majuscules). Il est suivi d'une tabulation. <i>Attention : c'est le seul endroit dans une transcription ICOR où l'usage de la tabulation est admis.</i>	COR salut
1b	Participant incertain	L'identifiant supposé est suivi d'un point d'interrogation <i>Attention : dans le travail sous Clan ou d'autres logiciels d'alignement, l'utilisation du « ? » n'est pas acceptée.</i>	COR? salut
1c	Hésitation sur les participants	Porter en les séparant par un « ? » les identifiants des participants entre lesquels on hésite. <i>Attention : dans le travail sous Clan ou d'autres logiciels d'alignement, l'utilisation du « ? » n'est pas acceptée.</i>	SAR?COR salut
1d	Participant non identifié	L'identifiant est composé de trois « X », ou de deux « X » suivis d'un « F » ou d'un « H » selon l'identification du sexe du participant.	XXX salut XXF salut XXH salut
1e	Événement non attribuable à un participant	Description entre doubles parenthèses sans identifiant.	((un coup de tonnerre retentit))
2- Tour de parole. Le tour est une unité théorique qui correspond à la production verbale d'un locuteur (pour le concept de tour, cf. par ex. l'article Sacks, Schegloff & Jefferson (1974).			
2a	Notation du tour	La notation du tour est insérée après l'identifiant du participant et la tabulation correspondante et constitue un paragraphe. Deux cas de figure : - si un tour s'étend sur plusieurs lignes, le pavé de transcription du tour doit être aligné à l'aide d'une indentation (retrait de 1 ^{ère} ligne). L'usage du retour forcé à la ligne (avec ¶) suivi d'une tabulation est interdit.	FAC bonne journée madame CL1 oui vous aussi= FAC =merci au revoir CL1 au revoir JEB alors est-ce que ça veut dire TOI tu fais quarante pour cent du chiffre d'affaire sur tes produits qu` ont moins d` cinq ans/ ou est-ce que ça veut dire qu'y a eu une époque où effectivement t` as eu tu as eu des S:AUTS (0.3) inno- d'innovation forts/

		- si la représentation d'un tour s'étend sur plusieurs lignes et qu'une ligne d'un autre locuteur s'insère (chevauchement ou commentaire), il faut insérer le symbole « & » à la fin de la première ligne (sans espace avant) et au début de la reprise du tour (sans espace après).	CHA LAU CHA LAU	et on avait ici cette grande salle euh: .h euh qu'on avait appelé finalement biblio[thèque& [archives&& &archi:ves::] &&bibliothèque en fait]
2b	Enchaînement immédiat (ou latching : pas de micro-pause entre deux tours de parole)	Insertion du symbole « = » à la fin du premier tour (sans espace avant) et au début du suivant (sans espace après).	BEA CL6 CL5	orange/ j'ai\ (.) c'est bon/ zig zag orange/= =oui oui [voilà] [et av]ec ça:: (.) des gauloises °blondes s'il vous plaît°
2c	Chevauchement	Insertion de crochets « [» et «] », encadrant le chevauchement dans chaque tour. Les crochets ouvrant « [» (début du chevauchement) sont obligatoires (sans espace après) ; les crochets fermants «] » (fin du chevauchement) par contre sont facultatifs. Les crochets sont alignés verticalement au moyen d'espaces (attention : ne pas utiliser de tabulation).	CHA MAR LAU MAR CHA LAU	ça c'était un petit [peu l-] [ah ouais] donc attends (.) donc ça ce- cette façade a été complètement re[maniée] [voilà] c'est l` bâtiment en fait [dans: qu'on] [ah oui ça c'est:] ça [c'est la] cour intérieure ça\ [qu'on r`prend]
3- Actions				
3a	Action à valeur de tour attribuée à un participant	Description de l'action concernée entre doubles parenthèses.	FH FH	((frappe à la porte)) (0.7) bonjour/
3b	Action non attribuable à un participant	Description entre doubles parenthèses		((sonnette de l'interphone))

4- Structures segmentales (transcrire le verbal)

4a	Orthographe	L'orthographe française est utilisée pour transcrire le verbal Lorsque les caractéristiques phoniques le requièrent, l'orthographe peut être adaptée.	ELI voui (.) alors elle va: elle elle elle revient pas SI/ (0.4) SUZ SI mais: j` sais pas quand: peut-êt` euh (0.5) euh::: ou di- ou dimanche ou ELI ah voui ce:[: cette semaine/ SUZ [hm hm]
4b	Elision Elision standard Elision non standard : absence/non prononciation d'un ou plusieurs son(s) en général vocalique(s)	Est notée par l'apostrophe « ' » L'élision non standard est notée par une antiquote « ` » (altgr+7+espace sur PC) (touche dédiée sur Mac). Si l'élision non standard se produit en fin de mot, faire suivre l'antiquote d'un espace. <i>Attention</i> : le « e » muet en fin de mot n'est pas une élision non standard, et ne doit pas être noté comme élidé.	FAB de l'interlocuteur distributeur/ (0.3) en disant hein hein moi j'ai j'ai dix nouveautés là qui arrivent ELI qu'est-ce que: euh j- un p'tit gâteau: nan/ JEB arrête en fait même ça peut êt le jeu d la distribution/ c'est qu tu laisses la marque lancée/ (0.3) pendant six mois il soutient à fond tu sais qu ça va s v[endre pa`ce que] t as l phénomène/ SOP [ouais ouais] JEB pourquoi pa`ce que t` as quarante pour cent du chiffre d'affaire/
4c	Allongement	Le son allongé est noté par des « : » en respectant l'orthographe (attention : pas d'espace avant « : »). Les « : » sont répétés en fonction de la durée perçue de l'allongement.	ELI voui (.) alors elle va: elle elle elle revient pas SI/ JEB ou c`ui qui s` vend pas trop pa`ce qu'i`s vont dire ben faut bien::: (0.2) faut faire dix centimèt` là dans l` rayon
4d	Amorce de mot	Insertion de « - » après le son tronqué (attention : pas d'espace avant).	COR je retr- retrouve
4e	Liaison facultative faite (notation facultative)	Mettre un « ^ » entre les deux mots liés (attention : pas d'espace ni avant ni après).	COR on va être trop en avance
4f	Liaison obligatoire entre deux mots non faite	Mettre une barre verticale (pipe) « » avant le deuxième mot (attention : un espace avant mais pas après la barre) Cette barre se fait avec Altgr+6 sur PC ou Shift+alt+L sur Mac.	SAR les elastiques
4g	Nombres	Les nombres ne sont pas notés par des chiffres mais en toutes lettres (cf règles de l'écriture des nombres).	BEA soixante-dix cinquante\ (0.7) soixante-trei:ze/ (0.7) soixante-quin:ze/ (0.8) quatre-vingts/

4h	Acronymes, sigles	Un acronyme (sigle prononcé comme un mot) est écrit en minuscules, sans ponctuation ni espace. Un sigle (mot épilé) est noté en minuscules, sans ponctuation, séparées par des espaces.	LAU JEB	parce que déjà ça veut dire quoi on va aller se battre en utilisant l' onu etcetera et tu crées la différence avec l` m d d
4i	Noms propres	Ne pas mettre de majuscule à l'initiale des noms propres (noms de personnes, d'institutions, de lieux/toponymes, etc).	SIM	les autres j'ai pas eu de faire part encore sylvie euh a dit euh a::h je me marie je me marie au mois de mars
4j	Onomatopées	Voir liste en fin de tableau.		
4k	Segments inaudibles, dont le nombre de syllabes est identifiable	Les segments incompréhensibles sont représentées au moyen d'une série de caractères x, chaque caractère x ayant la valeur d'une syllabe.	CL1 FAF	c` qu'on fait xxx entre voisin [mais là] [oui/]
4l	Segments inaudibles, dont le nombre de syllabes n'est pas identifiable.	Les segments complètement inaudibles sont représentées par « (inaud.) ».	MAM PAP	oh ben ça m'étonnerait qu` ton père [(inaud.)] [(inaud.)]
4m	Multitranscriptions d'un même segment et segments incertains	Les segments entre lesquels hésite le transcripateur sont séparés par un point-virgule et notés entre parenthèses (attention : pas d'espace avant « ; »). Les segments incertains sont notés entre parenthèses simples.	PAP SIM	ah ben ça elles sont toutes passées (là-haut; là euh:) les autres =j'avais plus de responsabilité que quand j'étais (élève) avec vous
5- Pauses				
5a	Micro pause	les pauses d'une durée inférieure à 0,2 secondes sont notées par (.)	VIC	LÀ j'ai un rapport à FAIRE/ à RENDRE/ (.) dans: le: cadre/ de mon stage/
5b	Pause chronométrée	Les silences sont chronométrés à l'aide d'un logiciel au 10 ^{ième} de seconde , et notées entre parenthèses. La même convention est utilisée pour les pauses à l'intérieur du tour d'un locuteur et pour les pauses entre les tours de deux locuteurs successifs ; dans ce dernier cas les pauses sont notées sans identifiant de tour.	CHA CHA LAU MAR	=hein/ c'est ça\ .h mais pour l'instant on est pas: encore très finement (0.4) dans les coupes et autres/ ben ils pourront descendre par l'escalier/ en maillot de bain (0.3) non:: ils vont descendre non:
6- Productions vocales				
6a	Productions vocales : rires, pleurs, lecture, imitation d'un acteur ou d'une prononciation particulière, etc	La description d'une production vocale est notée entre doubles parenthèses et précède la transcription, l'ensemble est compris entre chevrons pour en indiquer la portée : <((DESCRIPTION)) TRANSCRIPTION>	IRM SIM	((rire)) et: euh (0.5) tu fais: ça (0.5) toujours le temps/ <((en bâillant)) ouais ouais ouais> (0.3) j'espère/

		Note : la notation d'un enchaînement direct (2b), ou celle d'un chevauchement (2c), précède la description de la production vocale.	COR je ne l'ai pas vue hier= SAR =<((en riant)) non::>
			COR je ne l'ai pas [vue hier SAR [<((en riant)) non::>]
6b	Aspiration (facultatif)	L'aspiration est notée par la lettre « h » précédée d'un point (attention : le point n'est pas suivi d'un espace).	CHA l'accès principal l'accès d` service (.) .h:: la seule chose (0.2) qui serait à modifier pour l'instant/ .h c'est on sait qu'y a pas de p- (.) de piscine extérieure\
6c	Expiration, soupir (facultatif)	L'expiration est notée par la lettre « h » seule.	MAR t` as pas la place de: CHA hm h:::

7- Prosodie

7a	Montée et chute intonative	Les montées et chutes intonatives sont notées par « / » et « \ » sans espace avant. Les montées et chutes fortes sont notées « // » et « \\ ». Note : la notation d'une montée ou d'une chute suit celle de l'allongement.	MAR oui on n'en- on les voit pas CHA on les voit pas\ CLA vous allez bien// CHA .h tu gardes toujours ton mur ici:/ (0.3) .h t` avais ta porte là:/
7b	Saillance perceptuelle	Les segments caractérisés par une saillance perceptuelle particulière (intensité accrue et autres) sont notés en majuscules. Note : Les majuscules doivent être accentuées quand il y a lieu.	MAN mais c'est dans mais c'est sor[ti de mon morceau] PAT [ouais c'est DANS LE] PAIN/ ÇA\ MAN mais c'est DANS l` PAIN
7c	Baisse du volume de la voix	Les segments où le volume de la voix est plus bas (chuchoté ou murmuré) sont notés entre ° °	JEB alors qu'i` s` disent (0.2) °c'est quand même bizarre/°

8- Descriptions

8a	Généralités	Les commentaires sont notés dans un paragraphe propre sans identifiant de tour (voir 1/3), avec tabulation, entre doubles parenthèses et précédé de « COM : ».	((fin de réunion))
8b	Passages non transcrits	Si la transcription de l'enregistrement est interrompue, on le note entre double parenthèse en indiquant le timing de début et le timing de fin du passage non transcrit.	((00:38:41)) ((passage non transcrit)) ((00:54:22))
8c	Alternance de code (notation facultative)	Si vous souhaitez noter ce phénomène, les mots ou segments en langue étrangère sont transcrits entre double parenthèse avec la portée du phénomène notée par des chevrons.	ISA <((en anglais)) i go to school>

9- Liste des onomatopées, interjections et autres particules (liste non exhaustive)

ah

aïe

bah

bè

ben

eh

enfin

euh

gnagnagna

hein

hum (raclement de gorge)

hm hm

moui

mouais

m` enfin

oh

oh la la

ouais

ouille

oula, ouh là

p

pf, pff

pt

.tsk (bruit d'ouverture de bouche)

wouf, waf

ye