

## ANR-06-CORP-006

### Échange de corpus d'apprentissage multimodaux (MULCE)



## Tâche Tmult1 : Multimodal, modèles, transcription

### Annexe technique : Mode d'emploi pour transcription multimodale adapté pour le projet Infral



Version du 20 aout 2009

#### Auteurs

Thierry CHANIER, d'après Maud CIEKANSKI et Anna VETTER, Projet CoPéAs, 2005

## 1. Table des matières

1. Table des matières .....	2
2. Mode d'emploi pour transcrire .....	3
2.1. Description de la trame transcription .....	3
2.2. Données primaires à parti desquelles se fait la transcription .....	3
2.2.1. Listage .....	3
2.2.2. Organisation des données .....	4
2.3. Organisation générale feuille "session" .....	5
2.4. Modalité clavardage .....	6
2.5. Modalité production .....	7
2.6. La modalité audio et la table trancription oral .....	8
2.7. Anonymisation .....	11

## 2. Mode d'emploi pour transcrire

Concernant le sens des mots utilisés ici sur les interactions multimodales en ligne (mode, modalité, session, séquence, tour de parole, silence, etc.), on se reportera à :

Chanier, T., Vetter. A. (2006) "Multimodalité et expression en langue étrangère dans une plate-forme audio-synchrone". *Apprentissage des langues et Système d'Information et de Communication (Alsic)*, vol. 9. pp 61-101. [http://alsic.u-strasbg.fr/v09/chanier/alsic\\_v09\\_08-rec3.pdf](http://alsic.u-strasbg.fr/v09/chanier/alsic_v09_08-rec3.pdf)

### 2.1. Description de la trame transcription

Le fichier Excel dans lequel le transcripteur va effectuer la saisie des données se compose des feuilles suivantes:

- une feuille "liste session": elle comporte des informations sur les sessions, les documents correspondants et les transcripteurs.

Le transcripteur complètera cette feuille en reportant pour chaque session transcrite les informations afférentes. **Ces informations sont très importantes. Il faut donc les vérifier avec beaucoup d'attention.**

- une feuille "acteurs": cette feuille comporte notamment des informations sur les acteurs nécessaires à la transcription (3 premières colonnes). Elle ne doit pas être modifiée.
- une feuille "codes": elle comprend l'ensemble des valeurs que le transcripteur copiera et collera, de façon à éviter les erreurs de frappe. Ces valeurs sont organisées par catégorie (actions parole, actions production, actions sur les objets) et on y retrouve l'ensemble des codes d'anonymisation pour l'ensemble des acteurs du corpus.
- une feuille par "session" qui porte le code de la session transcrite (ex : `infral_q2_1`). C'est dans ce gabarit que le transcripteur saisie les données. Il ajoute une nouvelle feuille par session supplémentaire transcrite.

Le fichier " `Infral_transcript_patron.xls`" est le patron fournit au départ. Lors des transcriptions on veillera à changer le nom en indiquant en fin de fichier la date du jour et les initiales du transcripteur (pour éviter confusion entre plusieurs versions). Par exemple, "`Infral-transcription_evh_090821.xls`" pour le travail de transcription d'une ou plusieurs séquences faites jusqu'à ce jour par Eva-Maria H. et . "`Infral-transcription_evh_090822.xls`" pour la poursuite du travail le jour suivant.

### 2.2. Données primaires à parti desquelles se fait la transcription

#### 2.2.1. Listage

Pour chaque session un ensemble de données dites "primaires" vont servir de base à partir de laquelle la transcription sera faite.

Il s'agit en particulier :

- Fichier vidéo de la session
- Fichier trace du clavardage
- Autres fichiers, comme : fichiers correspondant aux images extraites des diapos montrés lors des sessions Centra.

Une partie de ces données sont des données primaires au sens où ce sont recueillies lors de l'expérimentation et contenu dans le dévédérom fournit à chaque chercheur.

D'autres peuvent être de nouveaux fichiers, par exemple, si le transcripteur décide d'extraire des images de la vidéo correspondant aux diapos montrées dans la session Centra. On fera alors attention de bien nommer ces fichiers avec le nom figurant dans la partie Agenda de la session Centra.

Il peut également s'agir d'une nouvelle version de la vidéo. En effet, le format vidéo primaire est celui de Centra (exemple : `infral_q2_108-11-03_12.11.exe`). Il est très compact, mais peu pratique pour la transcription car il ne permet pas d'accéder directement à une partie de la vidéo à partir d'un curseur temporel, ni de ralentir le déroulement de la vidéo. Pour voir la différence, un fichier de type AVI est fourni à titre d'exemple (`essai_infral_avi_q2_1.avi`). Il s'agit d'un extrait de la vidéo Centra. Nous conseillons de le lire à l'aide d'un programme gratuit tel VLC pour voir la différence. VLC est un bon outil pour transcrire les vidéogrammes. Il appartient au transcripteur ou à son équipe de faire les traitements nécessaires pour passer du format Centra au format AVI, s'ils en estiment le besoin.

Ce travail serait aussi nécessaire, car il n'est pas raisonnable de stocker des données vidéos dans un corpus dans un format propriétaire (Centra) au lieu de ouvert (AVI).

### **2.2.2. Organisation des données**

Besançon a fourni un dévédérom à chacun des chercheurs contenant les données primaires recueillies lors de l'expérimentation.

Si lors de la transcription de nouvelles données sont produites, dérivées des données primaires (vidéos aux formats différents, images extraites, etc.), alors il appartient de mettre l'ensemble (données primaires, données dérivées, dans un même dossier dont le nom sera l'identifiant de la session.

Par exemple pour la session 1 du quadrem 2, cela pourrait donner :

Dossier "`infral_q2_1`" contenant :

- `infral_q2_108-11-03_12.11.exe` (vidéo Centra, fournie dans DVD)
- `centra1_q2 - clavardage.html` (clavardage fourni dans DVD)
- `infral_q2_1.avi`, vidéo au format AVI ( non fournie)
- `cartemagalie.jpg`, image diapo utilisée pendant session (non fournie)

Ces dossiers seraient ensuite transmis à tout le groupe de recherche et partagé et viendrait ainsi compléter les données fournies par Besançon.

### 2.3. Organisation générale feuille "session"

Chaque feuille de transcription se présente ainsi : une boîte de renseignement, rappelant en particulier l'identifiant de la session et le nom du fichier vidéo de la session.

vidéo : infral\_q2\_108-11-03\_12.11.exe  
 id session : infral\_q2\_1  
 le temps de la vidéo qui commence à 0 est celui qui figure dans le haut de l'écran Centra  
 les commentaires de cellule (pas la colonne commentaires) sont pour Eva, donc à retirer ensuite

id_sequence	id_action	Canal	Début	Fin	Code_Acteur	Langue	Contenu
-------------	-----------	-------	-------	-----	-------------	--------	---------

Liste des colonnes :

- .id\_sequence : une session est composée d'une suite de séquences (voir (Chanier & Vette, 2006)) que le transcripteurs délimitera lors de sa transcription. Il donnera à chacune un identifiant unique, permettant donc de la distinguer de toutes les autres dans une même feuille. (donc deux séquences peuvent avoir le même identifiant dans deux session différentes).
- id\_action : c'est l'identifiant d'une action correspondant à un canal donné. Chaque action (une ligne par action dans la feuille Excel) a donc un identifiant différent.
- Canal : le canal ou modalité dans laquelle s'est produite l'action, soit
  - tpa : tour de parole audio
  - tpc : tour de parole clavardage
  - v : vote (cadre "participants" dans Centra, composant "communication dans (Chanier & Vetter, 2006))
  - prod : action dans le cadre outils de production collaborative dans Centra (peut-être un tableau blanc, un écran partagé, une diapo, etc.)
  - abs : icône d'absence momentanée (cadre "participants" dans Centra, composant "communication dans (Chanier & Vetter, 2006))
  - lm : main levée pour demander la parole (cadre "participants" dans Centra, composant "communication dans (Chanier & Vetter, 2006))
- Début, Fin : temps de début et fin d'une action repérée **uniquement** par rapport au temps de la vidéo tel qu'indiqué dans l'écran Centra, en haut et à droite de la fenêtre correspondante.
- Code\_acteur : cf. feuille acteurs où figurent ces codes
- Langues, langue majoritaire du contenu de l'action. Les langues sont codées suivant ISO-639-3 à 3 caractères. Chaque cellule de cette colonne doit avoir un contenu (pas de vide), donc on utilisera un code spécial si la langue ne peut s'appliquer à l'acte. Donc, dans Infral
  - deu : allemand ;
  - fra : français ;

- na : non applicable
- Contenu de la transcription : cf. Tableau 1.
- Commentaire\_transcripteur

<i>canal modalité</i>	<i>valeur</i>	<i>Début, Fin</i>	<i>Description</i>
abs	away back	Temps début = temps fin (donc action instantanée)	Transcription de l'icône sortie momentanée au moment de son apparition à l'écran  Icône de sortie momentanée disparaît (acteur revenu)
es	enter leave	Temps début = temps fin (donc action instantanée)	Arrivée d'un acteur dans Centra  Acteur quitte session Centra
tpa	Voir section spécifique	Temps début = début intervention  Temps fin = fin intervention	Audio et silences
tpc	Voir section spécifique	Temps début = temps fin (donc action instantanée)	Clavardage
v	yes no	Pour chaque valeur, temps début = apparition icône à l'écran, temps fin = temps disparition icône (si icône reste longtemps à l'écran) ou temps début	Vote : tick vert  Vote : croix rouge
lm	up down	Temps début = temps fin (donc action instantanée)	Visualisation d'un lever la main  Fin visualisation d'un lever la main
prod	Voir section spécifique	Temps début = début action  Temps fin = fin action	Transcription des actions dans l'espace de production collaboratif.

Tableau 1

## 2.4. Modalité clavardage

Les contenus de tpc sont tirés directement des fichiers de clavardage, quand ils existent. Si, lors des transferts, les accents ont été modifiés (accents remplacés par signes bizarres), on veillera à les rétablir. Mais il ne faut en aucun cas, réécrire les contenus (correction, changement casse, etc.).

On notera cependant deux cas particuliers d'intervention :

(1)

[\_a : afbes2\_3] oublie ce que j'ai dit :-)

Il s'agit d'un tour de parole spécifiquement adressé à un interlocuteur (les autres ne le voient pas). On note alors en début de champ Contenu par [\_a : <code\_acteur>] puis le contenu tapé.

(2)

```
salut <acd cd="afbre2_4">annika</acd>
```

Dans le contenu tapé par un acteur, celui-ci s'adressait ou faisait référence à un autre acteur en l'apprenant par son prénom. On ajoute alors des blaises à la XML, permettant de repérer de qui il s'agit en indiquant le code de l'acteur. On se garde bien de rectifier ce qui a été tapé au départ. Ainsi, on ne remplace par "annika" par "Annika". De même si le prénom n'est pas orthographié correctement on laisse ainsi.

Dans certains fichiers clavardage, les temps des tours de parole ne figurent pas. On essaiera alors de les insérer au mieux lors de la transcription, en sachant que l'incertitude peut-être grande. Mais ces interventions sont souvent importantes pour comprendre le dialogue d'ensemble.

## 2.5. Modalité production

Les actes production correspondent aux actions survenant dans les deux cadres de l'écran Centra, celui noté "Agenda" ou le principal où apparaissent les diapositives, le tableau blanc, etc. Il pourrait éventuellement s'agir d'une fenêtre indépendante partagée (quand un utilisateur ouvre un programme sur son ordinateur et le partage avec les autres).

Au contraire de l'environnement *Lyceum* (Chanier & vetter, 2006), les actions des outils de production collaboratives n'ont pas été analysées ni répertoriées, donc pas codifiées. Le transcripteur est donc libre en sachant qu'il doit mettre un minimum d'information, de façon la plus cohérente, systématique possibles en vue de permettre plus tard à d'autres transcripteurs de les reprendre.

Quelques exemples :

(3) diapo("ce qui va se passer")

une diapo apparaît à l'écran, on reproduit le nom tel qu'il figure dans l'agenda

(4) agenda(chargement, "Karte presentation.bmp")

l'acteur vient de charger dans la partie agenda une diapo qui n'est encore pas visible à l'écran

(5)

plus de diapo, écran de base de Centra

Si on décide de transcrire le contenu d'une diapos (la partie texte du moins), on pourrait alors imaginer le mettre en argument supplémentaire

(6)

diapo("ce qui va se passer", "ce qui va se passer, présentation de petits codes pour que l'on puisse respecter le temps de parole de chacun ...")

ou

(7)

diapo(ce\_qui\_va\_se\_passer.jpg, "ce qui va se passer, présentation de petits codes pour que l'on puisse respecter le temps de parole de chacun ...")

si l'on a réalisé une image JPG de la diapo, on met alors le nom du fichier correspondant.

## 2.6. La modalité audio ou table transcription oral

Un principe fondamental dans la transcription est d'assurer la continuité de la transcription de la modalité audio, sans interruption (Chanier & Vetter, 2006). Autrement dit, la suite des actes audio (tpa) doivent couvrir, sans trou, en continuité, tout l'espace temporel entre le temps de début et de fin de la session. Il se peut que par moments (qui peuvent être longs), aucun signal sonore ne passe dans la modalité audio (par exemple, quand les acteurs travaillent en silence dans le tableau blanc). On introduit un acteur fictif de code "sil" (pour silence) et le contenu de l'action correspond à la durée de ce silence. Le temps de début est le temps de début du silence et celui de fin, celui de la prise de parole suivante par un acteur. La différence entre les deux est la durée du silence.

On distinguera silence et pause (voir tableau ci-après). Les pauses durent, par convention jusqu' 3 secondes. Au-delà on ne note plus la pause dans une intervention d'un acteur, mais on introduit un acteur tpa de code "sil".

Les conventions spécifiques pour la transcription audio ont été reprises de la convention ICOR<sup>1</sup> et adaptée aux besoins de Mulce.

Conventions	symboles	Exemple de transcription
L'intonation montante est signalée par un slash /	/	ah /
L'allongement (insistance) est signalé par une mise en majuscule suivie de deux points (ajout du 7/07/07, d'après ICOR)	:	bonJOUR: BON:jour
La troncation est signalé par l'insertion de * après le son tronqué (attention : pas d'espace avant, un espace après)	*	thè*
L'élision non standard est notée par une antiquote.  Si l'élision non standard se produit en fin de mot, faire suivre l'antiquote d'un espace	`	il nous faut d`jà ça j` veux (un espace entre le signe d'élision et le mot "veux")
Les chevauchements en début ou en fin de tour de parole sont indiqués par les parenthèses	()	tpa1 : (Après) tpa2 : (ah) oui
Les chevauchements à l'intérieur d'un tour de parole sont indiqués à l'intérieur du tour de parole entre (<>) en indiquant le code_acteur concerné avec la balise stp, séparé du texte par des guillemets.	(<stp...>)	so + this is the document and you can write on here + (<stp cd="AT6" > yes + but euh </stp> but you need to just) + hum + talk together really just discuss the English Club site
Les pauses sont signalées par le signe + (1s) ++ (2s) +++ (3s)	+, ++, +++	the world wide web ++ world wide it means international +++ is it right
Si la pause dure davantage, elle apparaît avec la durée notée en secondes dans une ligne spécifique codée tpa (acteur = sil)		7

<sup>1</sup> Groupe ICOR 2006, *La convention ICOR*, site CORINTE, <http://icar.univ-lyon2.fr/projets/corinte/>

Conventions	symboles	Exemple de transcription
Les passages marqués par des problèmes phonétiques qui ont une incidence dans l'interaction sont signalés dans le texte entre crochets	[ ]	so the [first] question
<del>La transcription phonétique de ces passages, à partir de l'alphabet phonétique international, apparaît dans la colonne Phonétique.</del>  NB : nous utilisons la police SilDoulos		{fɔrst}  Non utilisé dans Infral
L'alternance codique est signalée par les accolades.  Le transcripteur indique le passage d'une langue à l'autre en notant entre accolades la partie du message qui se réalise dans l'autre langue.  Le transcripteur reporte pour chaque tpa la langue dominante dans la colonne réservée à la langue (notation iso). La langue caractérise la langue dominante d'un tpa, et ne vaut que pour les tpa (sinon, code "na" pour non applicable).  Si une partie du message est entre accolades, alors cela signifie que l'autre langue mise entre accolades est minoritaire.  Sur silence ou, binette – smiley- ou passage court du genre "ok" (valable dans deux langues), mettre dans la colonne Langues "na" pour non applicable.	{ }	{ah oui}
Les passages inaudibles ou trop incertains pour être retranscrits	[_XXX]	yes [_XXX]
Les commentaires du transcripteur apparaissent dans la colonne « commentaires du transcripteur » de la table parole en référence au pointeur de commentaires dans le texte de l'audio. Le pointeur peut renvoyer à un moment entre deux mots ou peut s'étendre sur plusieurs mots.	[*] pour un pointeur dans le champ commentaire  [suite_de_mots] pour un passage du tour de parole	Dans champ Contenu : I will draw here [*]. could you help me    Dans champ Commentaire : AT5 traces 3 flèches depuis "to play"
Pas de majuscules dans le texte sauf noms propres (Google, Lyceum, Bill Gates...)		
	[_rires]	
	[_souple]	
	[_sinterrompt]	
	[_tousse]	
	[_chante]	
lorsqu'il n'y a pas de son associé au passage (différent de inaudible ou incertain), lorsque, par exemple, l'enregistrement a connu une interruption momentanée	[_vide]	

Tableau 2

Attention, les tours de parole transcrits dans les corpus relèvent de la communication didactique en ligne, dans un environnement spécifique. C'est donc ce contexte particulier d'interaction qui permet d'interpréter la communication transcrite.

### ➤ **Comment transcrire un tour de parole audio (tpa) ?**

Le tour de parole commence avec la prise de parole du premier locuteur et se termine quand le locuteur a fini de parler. Sa détermination ne se fait donc pas en se fondant sur l'icône "haut-parleur" de la plateforme (micro en jaune dans Centra).

On attribue les pauses au tour de parole précédent. Un tour de parole se termine quand un nouveau tour de parole commence. S'il y a 4 secondes de pause, alors on indique un silence. Il n'est pas possible d'avoir une pause suivi d'un silence, s'il y a plus de 4 secondes alors cela annule la pause, et on indique un silence.

En cas de tour de parole sans son, mettre [\_vide]. Si le tour de parole est inaudible, alors le mentionner avec [\_XXX].

### ➤ **Comment transcrire les problèmes audio**

#### *Si plusieurs interlocuteurs ont des problèmes audio, que transcrit-on ?*

- Cas 1 : une icône apparaît, il n'y a pas d'audio, alors on marque l'icône et le tour de parole vide ([\_vide]).
- Cas 2 : une icône apparaît, avec un soupir ou du bruit puis disparaît, alors on indique la présence de bruit dans la colonne commentaire et on transcrit le passage inaudible avec [\_XXX].
- Cas 3 : une icône apparaît puis bruit ou respiration puis prise de parole, alors on indique la présence de bruit dans la colonne commentaire, on transcrit le passage inaudible avec [\_XXX], puis on transcrit ce qui est dit.

#### *Quelle distinction faire entre inaudible et incertain ?*

Ce n'est pas la même chose pour le transcripteur. Si le canal audio est fortement perturbé, alors le passage est étiqueté comme inaudible [\_XXX] avec un commentaire dans la colonne commentaire de la table parole.

En cas d'incertitude, le passage incertain sera indiqué avec [\_XXX] avec un commentaire dans la colonne commentaire de la table parole. Les [\_XXX] ne tiennent pas compte du nombre de syllabes incertaines, ils indiquent uniquement une incertitude.

#### *Quelle distinction faire entre troncation et inaudible ?*

Il y a troncation si l'interruption est intentionnelle pour le locuteur, le passage est inaudible s'il y a perte du message.

### ➤ Les chevauchements

En règle générale, les chevauchements ont peu d'intérêt ici (situation particulière d'apprentissage).

On ne note pas le segment qui est chevauché, on indique simplement l'existence d'un chevauchement.

```
(8)
alors euh je me présente je m'appelle <acd cd="afbes2_3">Marjorie </acd> et je vais être
l'animatrice de ce euh de cette séance Centra donc {willkommen} ++ est-ce que c'est bon
pour tout le monde euh pour le clavardage / vous ai bon j` pense que vous avez tous vu
comment ça fonctionnait / ++ ( <stp cd="afbre2_4">oui ça marche</stp) vous pouvez [_XXX]
oui s'il vous plait /
```

dans l'intervention orale de l'acteur afbes2\_3 (info figurant dans la colonne "Code\_acteur"), il y a eu chevauchement en fin d'intervention par une intervention orale (audio) de l'acteur afbre2\_4 (noté avec la baise <stp>), qui a dit "oui ça marche". Ne pas confondre ce codage du chevauchement avec le moment où afbes2\_3 dit "Marjorie", ce que l'on code avec la balise "<acd>"

Par convention, on fait démarrer sur la même seconde le temps de début d'un tpa et la fin du précédent. Il n'y a pas pour autant chevauchement.

### ➤ Les commentaires

Dans les commentaires, on indique directement le code acteur (mais sans les balises). Les balises sont utilisées dans les tpa et tpc.

Les commentaires sont laissés à la discrétion du transcripteur.

## 2.7. Anonymisation

Suivant les licences *Creative Commons*, les indications concernant les individus doivent être masquées (transformées) sauf indications contraires (adresses, patronymes, etc.). Les prénoms peuvent être conservés.

Les acteurs sont identifiés à l'aide leur code, le tout mi entre balise XML. Si le patronyme figure, il est enlevé et substitué par le code.

```
(9)
Robin Tartempion n'est pas là.

devient

<acd
cd='TutR'><firstname>Robin</firstname><lastname>TutR</lastname></acd>
n'est pas là
(10)
Robin est là
```

devient

<acd cd='TutR'>Robin</acd> est là

(11)

je te donne mon adresse [thipanie.petit@gmail.com](mailto:thipanie.petit@gmail.com)

devient, par exemple,

je te donne mon adresse [thipanie.bidon@gmail.com](mailto:thipanie.bidon@gmail.com)