

# Manuel d’annotation en actes de dialogue pour le corpus Datcha

Nicholas Asher, Alexis Nasr, Robin Perrotin

May 2, 2017

## 1 Introduction

Ce manuel a pour objectif d’aider l’annotation des tours de dialogue par des étiquette d’actes de dialogue. De telles étiquettes indiquent la fonction d’un tour de dialogue au sein du dialogue complet. Il s’agit d’une annotation non relationnelle, elle n’établit pas de liens entre les tours de parole.

Les étiquettes sont relativement grossières, ceci afin de favoriser la vitesse de l’annotation manuelle ainsi que l’annotation inter-annotateur.

### 1.1 Jeu d’étiquettes

Le jeu d’étiquette est composé de neuf étiquettes simples et des deux opérateurs | et \* permettant de composer des étiquettes simples pour former des étiquettes complexes.

Les étiquettes simples sont les suivantes :

OPE	Opening
PRO	Problem description
INQ	Information question
CLQ	Clarification question
ACK	Acknowledgement
TMP	Temporal
STA	Statement
PPR	Plan proposal
CLO	Closing
OTHER	Other

Un étiquette porte généralement sur un tour complet, lorsqu’un tour de parole peut être associé à plus d’une étiquette, on combine les étiquettes simples à l’aide de l’opérateur |. Ainsi si le début du tour de parole correspond à l’étiquette A et la suite à l’étiquette B, on annote le tour à l’aide de l’étiquette A|B.

Souvent une étiquette est “majoritaire” au sein du tour, on l’indique à l’annotateur l’indiquer grace à l’opérateur \*. L’étiquette A\*|B par exemple, indique que le début du tour correspond à l’étiquette A et la suite à l’étiquette B et que l’étiquette A est majoritaire.

## 2 Etiquettes simples

These elementary dialogue act types are all properties of just one turn. They are not intended to denote relations to other turns or elements of turns. But they resemble dialogue act types in other annotation efforts, in particular DAMSL.

### 2.1 OPE opening

OPE décrit des tours de parole permettant d'initier la conversation. Il s'agit de tours de parole principalement formels mais qui peuvent néanmoins véhiculer de l'information, tel que le nom du locuteur.

**Exemples :**

OPE TC Bonjour M . \_CLIENT\_ Je m' appelle \_TC1\_ et je vais traiter  
votre demande .  
OPE CL Bonjour

### 2.2 PRO Problem description

PRO décrit le problème qui doit être résolu par le téléconseiller. Elle désigne en gros la réponse d'un client à la question *que puis-je pour vous*.

**Restrictions :**

PRO ne peut étiqueter que les tours du client.

**Exemples :**

INQ TC Que puis -je pour vous ?  
PRO CL Mon décodeur TV est tombé en panne hier . Plus aucun signe vie .  
Je souhaiterais connaitre la démarche à suivre pour le remplacer

### 2.3 INQ Information question

INQ étiquette une question requérant de l'information de la part du destinataire.

**Exemples :**

INQ Puis-je avoir votre numéro de téléphone?  
INQ Est-ce que votre décodeur est allumé?

**Difficultés :**

INQ peut être difficile à distinguer de CLQ

### 2.4 CLQ Clarification Question

CLQ étiquette une question portant sur de l'information déjà donnée par l'interlocuteur.

CLQ demande à ce dernier plus de précisions sur l'information donnée.

**Exemples :**

STA Votre décodeur a un probleme. Je vous invite le redemarrer dans 30 minutes.  
CLQ OK je redemarre le décodeur dans 30 mins?

**Difficultés :**

La frontière entre INQ et CLQ n'est pas toujours facile à tracer car certaines questions sont un mélange de clarification et de demande d'information nouvelle.

**Exemples :**

STA TC Votre décodeur a un probleme.

Je vous invite le redemarrer dans 30 minutes.

CLQ CL OK je redemarre le décodeur et le livebox dans 30 mins?

Nous avons utilisé ici l'étiquette CLQ car la question porte sur une action qui a déjà été évoquée précédemment, même si ce qui est demandé va au delà de ce qui a déjà été dit.

En revanche, dans l'exemple suivant, la question n'est pas étiquetée CLQ mais INQ car le message du client est clair mais trop général. TC n'a pas besoin de plus d'information sur ce qui a été dit, mais sur la nature de l'événement décrit.

STA CL j' ai un message d' erreur qui s' affiche quand je lance la TV  
INQ TC pouvez vous me l' indiquer ? INQ

## 2.5 ACK Acknowledgement

ACK décrit un acquiescement vis-à-vis d'un tour précédent, qui peut être une assertion ou une requête (souvent signalée par une expression telle que *je vous invite*).

### Exemples :

INQ TC Avez -vous branché le décodeur sur une autre prise murale ?  
ACK CL Absolument

Contrairement à ce que son nom laisse supposer, un ACK peut être négatif.

### Exemples :

STA TC Je vous invite a démarrer votre décodeur  
ACK CL Ce n'est pas possible

### Difficultés :

ACK peut être difficile à distinguer d'une réponse (voir la section STA).

## 2.6 TMP Temporal

TMP étiquette les tours de parole qui indiquent à l'interlocuteur une pause dans l'échange de messages, le temps, en général de procéder à une action.

### Exemples :

TMP TC Je vérifie une information , je vous remercie de  
bien vouloir patienter quelques instants .  
TMP TC Je vous demande un instant le temps de compléter  
votre dossier pour cet échange .

## 2.7 STA Statement

Toute affirmation qui n'entre pas dans les catégories PRO, PPR ou TMP. Par exemple la réponse à une question.

### Exemple :

INQ TC Pouvez vous me confirmer le nom du titulaire de la ligne fixe  
\_NUMTEL\_ , afin que je sois sûre d' avoir le bon dossier ?  
STA CL Mr \_CLIENT\_ FABRICE

**Remarque :** Les étiquettes INQ et STA sont utilisées pour contraindre des paires question-réponse.

## 2.8 PPR Plan proposal

PPR étiquette un tour qui propose une action, non encore réalisée, visant à résoudre le problème du client. Cette action peut être effectuée tout de suite (redémarrer une livebox ...) ou bien dans le futur (aller récupérer un nouveau décodeur, prendre un rendez vous téléphonique ...).

### Restrictions :

PPR ne peut étiqueter qu'un tour de parole du téléconseiller

### Exemples :

un PPR peut se prolonger sur plusieurs tours de paroles non contigus, comme dans l'exemple suivant où 25 est PPR car c'est la suite de 20.

```
PPR 20 TC Merci d' avoir patienter . Veuillez noter votre numéro d'
        échange afin de récupérer un nouveau décodeur ...
STA 21 TC Vous pouvez souscrire à l' option Enregistreur lors de
        l' échange , pour avoir le nouveau décodeur UHD 87 SAT .
INQ 22 CL Option incluse dans l' offre liveboxstar non ?
TMP 23 TC Je vérifie l' offre .
STA 24 TC Exact , l' option est incluse .
PPR 25 TC Donc à l' agence vous demander l' enregistreur et
        le nouveau décodeur UHD 87 SAT .
```

## 2.9 CLO closing

CLO décrit les tours de parole visant à clore une conversation

### Exemples :

```
CLO TC Orange vous remercie de votre confiance .
        Je vous souhaite une bonne journée .
CLO CL Bonne journée à vous également . Merci
```

## 2.10 NUL

L'étiquette NUL est utilisée pour tout tour de parole qui n'entre pas dans la problématique de la résolution du problème du client.