

# Repérage des métonymies des Entités Nommées en français : directives d'annotation

## Avant-propos

Les directives d'annotation suivantes sont une adaptation des guidelines élaborées par Malvina Nissim et Katja Markert dans le cadre de la tâche #8 “Metonymy resolution at SemEval-07” de la campagne d'évaluation SemEval-07.

La métonymie se définit comme “*a form of figurative speech, in which one expression is used to refer to the standard referent of a related one*”(Nissim et Market, 2007). La métonymie constitue donc une substitution, un raccourci permettant de désigner la localisation pour l'institution qui y est installée, l'organisation pour le produit qui y est manufacturé...

De manière analogue à Markert et Nissim, seules les classes sémantiques d'Entités Nommées **location** et **organisation** ont été repérées. Les catégories de chaque type de lecture à annoter sont ouvertes. Néanmoins, quelques régularités apparaissent, mettant en évidence des patrons métonymiques. Dans le cadre du repérage des métonymies des Entités Nommées pour le français, une sélection plus restreinte des catégories déterminées par Markert et Nissim a dû être effectuée. Nous précisons donc dans le présent document les patrons métonymiques retenus pour les deux classes sémantiques repérées (section 2), après avoir décrit le corpus de travail (section 1).

## 1 Collection des données

Le corpus de travail est une collection d'extraits issus du corpus InfoMagic “Côte d'Ivoire”, version 1.0, codé en XML. Ce corpus se compose de 14901 documents provenant de journaux quotidiens comme Apl, A.P., etc. Chaque extrait est appelé *Article* et comporte un nom de lieu ou d'organisation annoté pour la métonymie. Les deux sous-corpus contiennent respectivement 1000 articles pour les noms de lieu (`countriesFr.xml`) et 1000 articles pour les noms d'organisation (`companiesFr.xml`). Chaque article compte quatre phrases : deux phrases précédant la phrase contenant l'Entité Nommée à annoter,

la phrase dans laquelle le nom de lieu ou d'organisation apparaît et une phrase suivante. Il est à noter que dans le cas où le nom est présent au début ou à la fin du texte de base, l'article peut contenir moins de quatre phrases.

## 2 Directives pour l'annotation

Pour mener à bien le repérage des métonymies des Entités Nommées en français, nous avons utilisé l'analyseur syntaxique Xerox Incremental Parser (XIP, Aït et al., 2002) concernant l'extraction, ainsi que des lexiques locaux pour les annotations.

Les 2000 articles constituant le corpus de travail ont été sélectionnés parmi les 14901 documents du corpus "Côte d'Ivoire". Ces articles ont été annotés manuellement pour la métonymie. Pour chacune des occurrences appartenant à la classe sémantique `location` ou `org`, l'annotation métonymique a été effectuée en suivant des catégories spécifiques, précisées dans la section suivante. Seuls les articles dont la lecture a été jugée claire et sûre ont été retenus pour figurer dans le corpus de travail. De plus, le thème général du corpus (la Côte d'Ivoire) a imposé de lui-même une sélection des lectures : aucune allusion à un produit manufacturé en Côte d'Ivoire n'est présent dans le corpus, par exemple (donc la lecture `place-for-product` a été écartée). C'est pourquoi nous n'avons pas conservé l'intégralité des catégories définies par Nissim et Markert.

La proportion des lectures annotées contenues dans le corpus final suit la distribution réelle du corpus de base. Nous n'avons pas recherché à obtenir les mêmes proportions que celles présentes dans le corpus de SemEval-07.

### 2.1 Catégories de Lieux (*location*)

Concernant les catégories de lieux, seules les lectures littérales (`literal`) et métonymiques (`metonymic`) ont été retenues parmi les lectures définies dans les guidelines de semEval-07.

#### 2.1.1 Literal

La lecture `literal` pour les noms de lieux inclut une interprétation du lieu comme entité politique et comme locatif.

- (1) Les Norvégiens se trouvant déjà en **Côte d'Ivoire** devraient également envisager de quitter le pays. (\_00017)

- (2) Départ du président Charles Taylor pour le **Togo**. (\_00164)
- (3) "Qu'on range tous les moyens de violences pour maintenant essayer de remettre la **Côte d'Ivoire** sur les rails pour qu'elle redevienne le grand pays que nous avons connu", a lancé M. Wade. (\_01011)

## 2.1.2 Metonymic

Des patrons définis par Markert et Nissim, nous avons conservé les plus fréquents et ceux que nous avons jugé les plus sûrs, soit : `place-for-people` et `place-for-event`.

### x `place-for-people`

Le patron `place-for-people` représente les personnes, organisations associées au lieu. Il peut s'agir du gouvernement, d'une organisation – comme une équipe de sport –, ou bien de la population d'un pays.

- (4) La **France** annonce l'envoi de renforts pour assurer sa mission de contrôle de cessez-le-feu. (\_00666)
- (5) Les délégués devraient cependant se retrouver tous dans le même camp pour suivre la finale junior de la Coupe d'Afrique des Nations qui verra leur équipe nationale de football affronter l'**Egypte**. (\_03748)
- (6) La **Côte d'Ivoire** s'apprête à voter dimanche, sur fond de polémique électorale. (\_11929)

### x `place-for-event`

Le patron `place-for-event` désigne un évènement s'étant déroulé dans ce lieu ; une guerre civile par exemple.

- (7) J'ai fait l'**Allemagne** avec Siriki Konaté. (\_09050)
- (8) Le ministre des affaires étrangères français, Dominique de Villepin, pourrait l'évoquer dès mercredi avec le secrétaire général de l'Onu, Kofi Annan, en marge du débat sur l'**Irak** au Conseil de sécurité. (\_09519)
- (9) Pierre Mazeaud, coordinateur français de la table-ronde sur la **Côte d'Ivoire** à Marcoussis (Essonne), a été reçu vendredi matin par le président Jacques Chirac pour lui rendre compte des négociations entre les parties ivoiriennes. (\_02616)

## 2.2 Catégories des Organisations (organisation)

Comme du côté des catégories de lieux, pour les catégories des organisations, seules les lectures `literal` et `metonymic` ont été retenues parmi les lectures définies dans les guidelines de semEval-07.

### 2.2.1 Literal

Dans le cadre du corpus "Côte d'Ivoire", la lecture `literal` pour les noms d'organisation décrit les références à l'organisation en tant qu'entité politique en général. Il s'agira, entre autres, de la description de la structure de l'organisation, mais aussi d'une association entre organisations. Sont incluses dans cette catégorie les références aux noms de radios.

- (10) Le premier secrétaire du **parti socialiste**, François Hollande, a dénoncé dimanche "l'échec total" de la diplomatie française en Côte d'Ivoire, qui a élaboré "un accord inapplicable".(\_00119)
- (11) Les chefs d'Etat de la **CEDEAO**, qui se sont réunis vendredi à Dakar, ont décidé de reprendre l'initiative pour tenter de sauver les accords de Marcoussis (France), signés le 24 janvier par les protagonistes de la crise ivoirienne. (\_00908)
- (12) Les troupes françaises en Côte d'Ivoire ont pour consigne "d'ouvrir le feu" sur "toute personne empêchant manifestement la réalisation" de leur mission, a déclaré dimanche sur **France Inter** le colonel Emmanuel Morin, chef de corps du 2ème régiment étranger de parachutistes (REP) envoyé en renfort dans le pays. (\_00543)

### 2.2.2 Metonymic

Parmi les patrons disponibles de SemEval-07, nous en avons retenu trois : `organisation-for-members`, `organisation-for-facility` et `place-for-event`.

- x `organisation-for-members`

Le patron `organisation-for-members` représente les membres de l'organisation. Entre autre exemple, un acte officiel ou une déclaration formulée au nom de l'organisation, la participation à une action de tous les membres de l'organisation.

- (13) La semaine dernière, le **ministère des Affaires étrangères** a conseillé aux Français "dont la présence n'était pas essentielle, de quitter la Côte d'Ivoire". (\_07458)
- (14) Il convient également d'indiquer qu'au cours de cette cérémonie, **UNIWAX** a procédé à

la remise des lots du concours de pronostics organisé dans le cadre de l'élection Miss Côte d'Ivoire 2002. (\_12317)

- (15) Après la visite du **Conseil de Sécurité** Tévoédjré remercie Gbagbo (L'Inter 05/07/2003). (\_12759)

*x* organisation-for-facility

Le patron organisation-for-facility réfère à l'équipement qui abrite l'organisation, à l'instar d'un siège.

- (16) Une cellule chargée d'informer de la situation de la communauté française en Côte d'Ivoire avait été mise en place au **Quai d'Orsay**, joignable au 01 45 50 34 60. (\_02863)

- (17) Laurent Gbagbo ne participera pas au déjeuner à l'**Elysée**. (\_03553)

- (18) La première passation de pouvoir entre un ministre de l'ancien gouvernement ivoirien et son successeur, membre de l'opposition, dans le nouveau gouvernement de réconciliation nationale, a eu lieu vendredi au **ministère de l'Agriculture**, a rapporté samedi la presse ivoirienne. (\_09583)

*x* object-for-name

Ce dernier patron, object-for-name, désigne quand à lui l'organisation en tant que chaîne de caractères, citant ou explicitant un acronyme souvent mis entre parenthèses.

- (19) Les troupes rebelles du Mouvement patriotique de Côte d'Ivoire (**MPCI**) ont "carte blanche pour passer à l'offensive". (\_00976)

- (20) Le pouvoir en Côte d'Ivoire doit revenir à la rébellion du MPCI (**Mouvement patriotique de Côte d'Ivoire**).(\_03896)

- (21) N'oublions pas que MPIGO signifie **Mouvement populaire ivoirien du Grand Ouest** et son entrée dans la rébellion avait pour motif officiel, la vengeance de la mort de Robert Guéi. (\_10744)