

TWITTER...

ET LES PRATIQUES D'ÉCRITURE PROFESSIONNELLES / PROFANES

IR Corpus. Groupe de travail nouv-com. - 31 mai 2012 -
Intervention de Béatrice Turpin
Université de Cergy-Pontoise. EA 1392 CRTF/LaSCoD

TWITTER : LES PRINCIPES

Réseau social avec une pratique d'écriture spécifique

140 caractères maximum (**tweets**), avec possibilité :

- d'hyperliens internes au réseau, pouvant renvoyer :
 - à une **thématique** : ce type d'hyperlien est signalé par un croisillon **#** ou **hashtag** suivi d'un mot clé;
 - à un « **profil** » : **@mention** introduite par @nom d'utilisateur.
- d'hyperliens externes : images, vidéos, sites, blogs, autres réseaux.

Twitter définit également ses conventions de participation :
« **abonnements** » (Tweets que l'on décide de suivre) +
« **abonnés** » ou « **suiveurs** » (qui nous suivent)

→ interactions en réseaux

TWITTER : LES PRINCIPES (SUITE...)

Emetteurs pluralisés : messages souvent redirigés (« retweetés »)

→ + **Twitter via SMS** : possibilité de recevoir des tweets par SMS ou d'en envoyer à sa liste d'abonnement (via un code par pays).

UN PROJET D'ÉTUDE DOUBLE : LES PRATIQUES D'ÉCRITURE / LES REPRÉSENTATIONS DE LA BANLIEUE

Etude en phase exploratoire pour Twitter : pratiques d'écriture professionnelles (politiques pour Julien Longhi, journalistiques pour Béatrice Turpin)



ETABLISSEMENT D'UN CORPUS DE TWEETS JOURNALISTIQUES

Etablissement du corpus à partir de la page « **Catégories** » de Twitter en sélectionnant les institutions, soit tweets :

- de journaux papier (*Le Monde, Le Figaro, Métro...*),
- de journaux internet (*Rue 89*),
- de radios (*France Info, France Inter...*),
- de chaînes télévisuelles (*Arrêt sur images, BFMTV...*)



The image shows a screenshot of two Twitter profiles. The first profile is for Mediapart (@mediapart), which is verified. Its bio reads: "Les tweets de Mediapart.fr : le journal d'information politique et générale, d'investigation, indépendant et participatif." The second profile is for France Inter (@f_inter), with the bio: "Fil officiel de France Inter, la radio publique généraliste de Radio France". Both profiles have a "Suivre" (Follow) button and a dropdown menu icon.

MÉTHODOLOGIE

● Collecte manuelle des données ?

→ Abondance des tweets :

Collecte des tweets sur un mois = environ 1500 tweets pour *Le Monde*. Ce même journal affiche un total de 62 597 tweets. On en dénombre par ailleurs 63 913 pour *Le Figaro*, 29 760 pour *France Info*... (au 31 mai 2012)

=> lenteur de l'ordinateur pour remonter le fil des tweets.

● Collecte avec un robot ?

→ **HTTrack**

On peut limiter la profondeur et la taille des téléchargements.

Inconvénient : recueil d'informations brutes, difficiles à traiter par manque de structuration.

→ **ScrapBook** ? (extension Mozilla). Permet également de récupérer les tweets en précisant la profondeur des liens. Possibilité de les nettoyer dans un éditeur de texte. On revient ensuite au traitement à la main.

Collecte qui permet d'avoir en mémoire-image les tweets susceptibles d'être effacés du site.

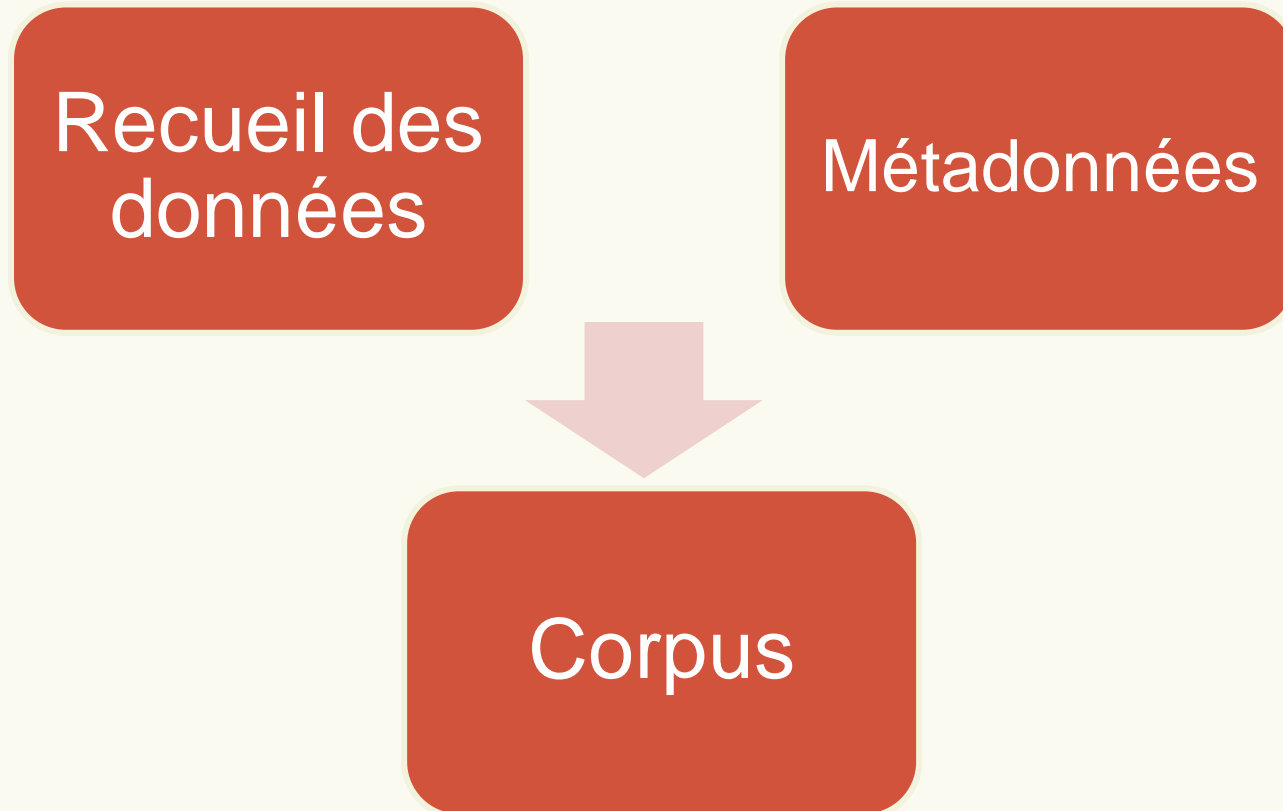
Nécessité de définir préalablement l'étendue du corpus.

MÉTHODOLOGIE (SUITE...)

● Etablissement des métadonnées

Métadonnées	Disponibles	Recherchées
Adresse de la page	X	
Auteur	X	X
Titre du média	X	X
Date et heure	X	X
@mention	X	X
Lien (codé; site <i>t.co</i> le décode)	X	X
Sujet (à partir du # mais pas seulement)		X
Etiquetage sémantique (ex. « sentiments »)		X
Etiquetage morphosyntaxique		X
Matériel utilisé pour poster		X
Autres ???		X

MÉTHODOLOGIE (SUITE...)



REMARQUES

→ Le choix des métadonnées guide/conditionne l'interprétation du corpus (ce sont les métadonnées qui font le corpus, qui est toujours déjà un interprété)

→ Distinctions :

- Analyse quantitative de corpus. Analyse de contenu ou TAL
- Analyse linguistique/textuelle qualitative de corpus. C'est dans cette dernière que s'insère le projet d'étude.

Twitter

- Support pour une analyse linguistique / textuelle qualitative
- Nécessité de faire une revue des études déjà effectuées sur le support (objectifs, thématiques, méthodologies, conclusions)

PERSPECTIVES COMPLÉMENTAIRES

Contraintes pesant sur l'écriture Twitter

Mesurer les facteurs de structuration de l'expression: type de discours, style de l'énonciateur, contrainte du support

- cf. Longhi, J. « Discours, style, format : contraintes et niveaux de structuration de la textualité des Tweets de Mouloud », *Actes du 3^{ème} Congrès Mondial de Linguistique Française* (à paraître)

Travail sur corpus hétérogènes: faire varier les types de discours (politique, médiatique), les genres de discours

- cf. Garric N. & Longhi J., éds, L'analyse de corpus face à l'hétérogénéité des données, *Langages* (en cours d'évaluation)

Hypothèses de sémantique discursive, rapports expression / 'contenu', etc.

- **Apports pour la linguistique de corpus, la sémantique, la linguistique du texte, etc.**