

Tutoriel Browser

Le principe

Chaque fichier contient des métadonnées au format IMDI. Le tutoriel des métadonnées est consultable [ici](#). Les métadonnées sont interrogeables à travers le formulaire de recherche qui se trouve dans la section "télécharger". Le formulaire permet de constituer des sous-corpus comprenant seulement les échantillons ayant des propriétés précises (par ex. tous les échantillons interactifs, ou tous les fichiers issus d'un certain corpus source).

Fichiers téléchargeables pour chaque échantillon

Une fois le sous-corpus obtenu, pour chaque échantillon il sera possible de télécharger les fichiers suivants:

1. Le fichier son (format WAV et MP3)
2. La transcription orthographique (format TXT)
3. L'annotation macrosyntaxique (format TXT)
4. L'annotation prosodique (format Praat et XML)
5. Les métadonnées (format IMDI et HTML)
6. Le pitch nettoyé (*-nett-pitch)
7. Le pitch lissé utilisé pour le calcul des contours globaux (*-liss-pitch)

Le formulaire

Recherche mots

Il s'agit d'un concordancier simple qui permet d'identifier les échantillons contenant un token ou une certaine suite de tokens.

Sex

Le sexe du locuteur, participant dans l'échantillon.

Valeurs :

- Indiff: non précisé
- Female: femme
- Male: homme

Age

L'âge du locuteur, participant dans l'échantillon.

Valeurs :

- Indiff: non précisé
- de 1 à 99 (son âge en chiffre)

Genre

Type de discours de l'échantillon

Valeurs :

- Indiff: non précisé
- Discourse : le contenu se compose d'énoncés produits dans le but de communiquer à son interlocuteur une information ou une intention.
- Ritual/religious text : Le contenu concerne l'accomplissement de rites religieux constitués de types de discours prescrits.

SubGenre

Type de séquence majoritairement représentée dans l'échantillon. Nous nous sommes inspirés, pour définir les valeurs de ce champ, de la théorie des séquences discursives d'Adam (2011), tout en gardant de la classification IMDI la définition du discours oratoire, qui est représentée dans notre corpus.

Valeurs :

- Indiff: non précisé
- Argumentation: construction, par le locuteur, d'une représentation discursive visant à modifier le point de vue de l'interlocuteur sur un objet de discours donné.
- Description : représentation des caractéristiques d'une personne, d'un objet ou d'un événement.
- Procedural : instruction, ou description des étapes impliquées dans l'exécution d'une tâche.
- Oratory : discours formel adressé à un public dans des contextes politiques, religieux, cérémoniaux ou légaux.

Corpus

Nom du projet source (pour une description des sources voir [ici](#))

Valeurs :

- Indiff: non précisé
- PFC : source externe
- CFPP2000 : source externe
- C-PROM : source externe
- ESLO : source externe
- Corpus Avanzi : source externe
- Corpus Lacheret : source externe

- Corpus Mertens : source externe
- Rhapsodie PROF : source interne
- Rhapsodie Movie : source interne
- Rhapsodie BROAD : source interne

Interactivity

Caractérise le degré d'interaction entre les locuteurs présents dans l'échantillon.

Valeurs :

- Indiff: non précisé
- Interactive : discours caractérisé par un échange entre les locuteurs
- Semi-Interactive : discours ponctué par des interjections répétées de la part de l'interlocuteur
- Non-Interactive : discours caractérisé par de longues séquences que le locuteur produit sans s'attendre à des réponses verbales de la part de son interlocuteur

Social Context

Indique le contexte situationnel au sein duquel l'événement prend place.

Valeurs :

- Indiff: non précisé
- Private: l'accès à la situation communicative est limité à un nombre restreint d'individus
- Public: l'accès à la situation communicative est ouvert à tous

Event Structure

Indique la structure monologale ou dialogale de l'événement langagier.

Valeurs :

- Indiff: non précisé
- Monologue: discours produit par un seul locuteur
- Dialogue: discours produit à plusieurs voix

Channel

Indique le canal de communication.

Valeurs :

- Indiff: non précisé

- Face-to-face : la transmission du message est centrée sur l'interaction multi-modale entre le locuteur et ses interlocuteurs
- Broadcasting: le message est transmis à un public large à travers un moyen de communication de masse (radio, télévision, etc.)

Planning type

Indique dans quelle mesure le consultant planifie l'événement linguistique.

Valeurs :

- Indiff: non précisé
- Spontaneous : discours sans aucune planification
- Semi-spontaneous: discours avec une planification partielle au préalable
- Planned : discours dont le sujet et la structure sont préparés à l'avance

Accès échantillon

Permet d'accéder directement à un échantillon de la base