Cordial : Utilisation du logiciel dans le cadre d'analyses linguistico-communicationnelles.

Nous vous conseillons la lecture du chapitre dans ouvrage Panckhurst (1999), « Analyse linguistique assistée par ordinateur du courriel », in Anis J. (dir.), *Internet, communication et langue française*, Paris : Hermès, p. 55-70, dans lequel nous exposons une utilisation du logiciel Cordial. Le but de cette section consiste à montrer comment on peut procéder à des analyses morpho-syntaxiques, à l'aide d'outils comme Cordial ou Nomino, dans des domaines divers. Dans l'article cité, il s'agissait d'effectuer des recherches sur l'utilisation du courrier électronique par des étudiants et des enseignants.

Cette section sera divisée en deux parties : la correction grammaticale ; l'étiquetage morpho-syntaxique.

Pour lancer le logiciel, cliquer sur l'alias « Cordial Analyseur », qui se trouve sur le bureau.

Avant de lancer la correction grammaticale dans Cordial, nous vous conseillons de vérifier les paramétrages dans le menu « Configuration ».

Plusieurs intercalaires permettent de régler la syntaxe (signalement de phrases incorrectes, phrases sans verbes, réglage de la sensibilité de la correction, etc.), l'orthographe et la typographie. Voyez l'intercalaire « Syntaxe : style » au sein du menu « Configuration/Paramétrer la correction ».

Plusieurs vérifications sont possibles pour l'orthographe :

Le dernier élément du menu « Orthographe » permet de retenir tous les mots du texte et le nombre d'occurrences. Ce travail peut être utile dans le cas d'un repérage de concordances.

Dans l'exemple ci-contre, la phrase contient une erreur lexicale « premie » ; on lance la vérification orthographique en cliquant sur :

La fenêtre proposant une correction de dictionnaire s'affiche ensuite.

Une fois que la correction est lancée, l'utilisateur peut intervenir de plusieurs façons, en cliquant sur les icônes qui se trouvent en bas de l'écran, et ce, à partir des explications qui apparaissent dans la zone de correction.





Prançais de France C Français du Canada	C Simaler les auto saves	Free contractions of
🗸 Simula las inconstitute de stale	Signale in not need the	Enregistrer s
Signaler Lagramuticalité d'une phrase	E Signaler les nots iniurieus	
Simale les désannes sémantiques	Simpler les mots renesiers	
Signaler les visules manquantes	E Signaler les mots anotiques	
Signales les formes passives	E Signaler les mots populaires	
Signaler les clichés	Signaler les mots familiers	
Signaler les alitérations	Signaler le tutoiement	
ANNONCER LES		
ANNONCER LES Iomophones: Suspects I Homoprophes: Suspec	ts 💌 Paranymes: Suspects 💌	
ANNONCER LES fomophones: Suspects I Homographes: Suspec SIGNALER I Angleismes I Conadamismes I Occiveri	te Y Paronymes: Suspects Y	
ANNONCER LES tomophone: Surpects Thomographes: Surpec SIGNALER PAggletimes Conagarismes Cocolaris Lebvétimes Atigarismes Régional	ts V Paronymes: Sutports V mes I Bylgicismes ames V Margues déporées	
ANNONCER LES tomophones: Surpects THomographes: Surpec SIGNALER TApplicitmes Consplayismes Constraints Behveitimes Abigarismes Regional OK Anno	ts Paronymes: Suspects Behjicismes mes Margues déposées det <u>60057055</u> <u>Aide</u>	





Après avoir réglé les paramétrages pour que l'analyse soit la plus fine possible, procédez à une vérification de phrases. Remarque : dans cette version du logiciel, qui est plus basée sur l'analyse morphosyntaxique que sur la correction, la possibilité d'effectuer la correction n'est pas toujours active.



Cordial propose une analyse morpho-syntaxique de la phrase. La version « Universités » de Cordial (qui est la version que nous utilisons) permet également plusieurs options supplémentaires : notamment l'étiquetage de texte et l'extraction de phrases sur occurrences et co-occurrences. Ces options se trouvent en bas du menu « Syntaxe » (« Étiquetage de texte », « Étiqueter un ensemble de fichiers », etc.) :

Pour l'étiquetage, le paramétrage par défaut est indiqué à la fenêtre ci-contre. Le choix de l'abrégé en majuscules pour le type grammatical (plutôt qu'un affichage numérique) permet de visualiser rapidement la catégorie syntaxique attribuée. On peut, bien entendu, maintenir ou désactiver les différentes options proposées en fonction de ce que l'on recherche.

À partir de la phrase « Voici un exemple d'étiquetage d'une phrase », nous avons fait un essai d'étiquetage. Après avoir introduit les réglages pour les paramètres, une fenêtre de dialogue classique, vous demandant le lieu de sauvegarde et le nom de votre fichier, s'affiche. Nous l'avons intitulé « exemple » (et son extension est .cnr ; le fichier est lisible dans Cordial, en demandant son ouverture depuis le menu « Fichier »). Les résultats apparaissent dans le fichier « exemple.cnr ».

Certains codages peuvent paraître flous dans un premier temps. Si on souhaite avoir uniquement accès aux simples catégories grammaticales, il suffit de décocher les différents options proposées dans la fenêtre de paramétrage. Dans cet exemple, nous n'avons retenu que la numérotation des mots de chaque phrase et le type grammatical en abrégé.

 Allichage de finitioducteur "	phrase : lo la mot C En fin de la	145E	 Ligne vide entre les phraces Ligne de titre en début de fichier
Numéroter les paragraphes Numéroter les phrases Expressions Aucure expressions	Marquege des dialogue Rolevé des ambiguite Locatione prono	N 🔽 Lemmex V Lemmes lén	Mot de codage spécieles w> masouin □ lagressans etimit airs
E Location percention	Encerniorn and		E Esperarion periodes
Type gramatical : C Augun C Nu	mérique 🤉 Abrégé e	n majuaculaa	
Codage spécialisé : Clausura Cla	fres @ Lettres +	espaces	
Appartemence is un groupe synta Fonction grammaticals Tratement des consums:	gnatique 🔲 Type 💌 🕅 Nund	de la proposition ro de la proposition	 Veibe de la proposition du moi diquivalente sémantiques
C Coniger et signaler les erreurs		C Coniger et n	e pas signaler les erreurs
C Ne pas corriger, signalar les en	80%	 Ne pas carri 	gar, na pas signalar
- Statistiques : V Antiguités Catégorise grammaticales V Personnes	 IF Northre des mote IF Genre des mote IF Temps verbaus 	I▼ Codages numé I▼ Fonctions gram I▼ Types d'advest	riques des lypes premiericaux (D à 201) molicalise les



 Jacketage de Patricitadeus Imme Alfadege da terminateur Imme Manifectuice der sons de chagar En debut de ligne C. App 	Processor C En in de lans C NON	Clighe vide entre les phrases R Lighe de title en debut de ficher
C Assesse for les paragraphes C Assesse for pleaser	T Marg age der diakog er 🗖 Lerviser T Ratevil der anbigsker 🗖 Lerviser för	I ⁿ Mot de codage spécialis Lo Nateulo
F durre openien F contre openien F contre openien	E anterapteriette E attente attendente E farmanischer vers	 Control of a feature of the second sec
Type germalical C Auton C N	nárious ⁶⁷ Adrápi or hekutovici	
Codage spicialisé 14 August C. La	tese Cuther + separar	
Appatenance kun gospe ore Foreien garmaitale Indenset dis conset: Conset of spisie les meau	egneliger 🔲 Type de la proposition 🗖 Numéro de la proposition C. Conjour et r	Velbe de la proposition du sont e quivalents rémaniques e (ex signifer les reneus
C Ne pas coriges, signaler les e	reun G Ne par con	ger, ne par sigraler
Catilopation Catilopation Personne	F. Nettlee des nets F. Conlagen runnl Geven des nets F. Fonctions give Traps verbaue F. Types d'advect	inaerden (zen ganandioaar)(? 3.221) instrales He

L'étiquetage proposé se lit de la manière suivante : voici (préposition), un (déterminant indéfini masculin singulier), exemple (nom commun masculin singulier), d' (préposition), étiquetage (nom commun masculin singulier), d' (préposition), une (déterminant indéfini féminin singulier), phrase (nom commun féminin singulier), « . » (ponctuation forte).

Pour des travaux en linguistique-informatique, notamment lorsqu'on souhaite travailler sur les résultats de fichiers étiquetés, il est souvent important d'avoir accès aux mots lemmatisés. Dans ce cas, à partir de la phrase « Voici un exemple d'étiquetage d'une phrase » contenant un codage lemmatisé, nous obtiendrons ce qui apparaît ci-contre.

Les occurrences des mots (ainsi que leurs catégories grammaticales) peuvent être consultées dans la fenêtre qui apparaît après avoir choisi « Syntaxe », « Occurrences des mots… »

La lemmatisation consiste à rapporter une unité à une forme canonique, ce qui aura pour effet de faciliter le travail ultérieur de la machine. Par exemple, un verbe fléchi sera généralement lemmatisé en verbe infinitif, un adjectif en masculin singulier, un nom en singulier, etc.

Dans l'exemple ci-contre, nous avons ajouté une phrase. Que remarquez-vous à propos de la lemmatisation ?

Travail à faire pendant la séance :

Procédez à l'étiquetage d'un texte. Vérifiez les catégories grammaticales proposées par Cordial pour votre texte.

Quelles différences voyez-vous entre « occurrences orthographiques » et « occurrences syntaxiques » ?

L'étiquetage est-il satisfaisant ? Quels sont les types de problèmes qui se posent ?

Piste : pensez aux différents problèmes qui se posent en traitement automatique des langues, notamment, l'ambiguïté.

Faites des essais au sein des configurations (choisir, au sein du menu « Configuration », l'élément « Paramétrer la correction », puis l'onglet « Typo. Séparateurs ») : quelles différences remarquez-vous au niveau des résultats des occurrences lorsque vous choisissez soit « Retourchariot est une fin de phrase », « Toujours », « Jamais », ou « Devant majuscule » ?

÷	exemple.cm	
N°	mot mot	Typegram
	==== DEBUT DE P	PHRASE ====
1	Voi ci	PREP
2	un	DETIMS
3	exemple	NCMS
4	d'	PREP
5	étiquetage	NCMS
6	ď	PREP
7	une	DETIFS
8	phrase	NCES
9		PCTFORTE
	===== FIN DE PH	RASE =====

🖤 exemple.	cni		
N° mot	mot	lemme	Typegram
====	DEBUT DE P	HRASE ====	
1	Voici	voici	PREP
2	un	un	DETIMS
3	exemple	exemple	NCMS
4	d'	de	PREP
5	étiquetage	étiquetage	NCMS
6	d'	de	PREP
7	une	un	DETIFS
8	phrase	phrase	NCES
9	contenant	contenir	VPARPRES
10	un	un	DETIMS
11	codage	codage	NCMS
12	lemmatisé	lemmatiser	VPARPMS
13	PCTF	FORTE	

===== FIN DE PHRASE =====

Uccurrences syntaxiqu	ies des mots			N	_ 0
ndane	NMS	1	-	Nombre de mots:	22
orteniz	VEDBE	1		Nombre d'occurrence	es: 19
1	PREP	2	_	Type des nots	Tous (spes
aes áticulatoria	NMS	÷		Morris 1	E Bonnes
axamnia	NNS	1		🕅 Acjediřa	M Advertices
exemple d'étiquetage	NMS	÷		M Affors	Pérostions
ermer	VERBE	1		Verres	E Corientena
1	PPER3S	1		F. Mair en meiurait	- E Chitras
mportant	AMS	1		T News commune	Nove croce
emmatiser	VERBE	1		1º Home contrigue	10 House propre
85	AUPIG	1		OK.	Coloriger
onrase	NE S	-		-	1
arobierne	DDED	1		Contegluel	Becheichei
-	HALF .	÷.			1
es bhrase sroblème sur	ADPIG NES NMS PREP	1 1 2 1 7	_	OK. Coniegtuel	<u> </u>