Manipulation de correcteurs orthographiques

(1) Etant donné le travail de repérage effectué à partir d'un corpus de phrases lors des deux séances précédentes, vous avez mis à jour les dysfonctionnements des logiciels selon :

- les fautes bien détectées, et bien corrigées,
- les fautes bien détectées, mais mal corrigées,
- les fautes signalées à tort,
- les fautes non détectées.

(2) A l'issu de ce premier travail, vous avez observé que certains phénomènes de langue ont été source de problèmes pour les correcteurs ; ils concernent par exemples :

- les problèmes lexicaux (noms propres, mots composés, ellisions …) ;
- le groupe nominal (accord déterminant-nom, adjectif, conjonctions …) ;
- le lien sujet verbe (accords sujet-verbe, adverbes, subordonnées …) ;
- le lien sujet attribut (accord sujet-adjectif, participe …) ;
- etc.

Ils s'expliquent généralement par des lacunes dans les entrées de dictionnaires, ou bien par l'absence de règles adéquates (ex. le placement des adverbes selon leurs propriétés lexicales ; l'inversion du sujet, les ellipses en comparatives), un défaut d'analyse structurelle (ex. où s'attache la relative ?), ou sont dus simplement à une trop longue distance entre deux termes pour le correcteur (ex. accord, antécédent de la relative, anaphore).

Devoir

Dossier de 5 à 10 pages, à rendre le jeudi 20 avril 2000, comportant votre corpus et le travail suivant :

Concentrez-vous sur deux ou trois des phénomènes problématiques rencontrés et organisez-les selon la catégorie de difficulté à laquelle ils appartiennent (voir (2)). Choisissez un seul logiciel.

Procédez par méthode :

(i) Notez et commentez précisément le comportement du correcteur ;

(ii) Expliquez ce comportement : faites des hypothèses sur les capacités et/ ou l'organisation des connaissances du logiciel face à l'analyse linguistique du problème de langue considéré (utilisez les ouvrages de grammaire ou de lexicologie pour faire cette analyse – quelles sont les règles ? –, et orientez votre corpus en conséquence – usages corrects vs. fautifs ou constructions agrammaticales à détecter vs. constructions grammaticales) ;

(iii) Proposez des solutions pour améliorer le logiciel et parvenir à traiter les phénomènes étudiés (n'oubliez pas qu'il lui faut trouver des indices dans la séquence linéaire pour pouvoir déclencher une règle).