

TRAITEMENT INFORMATIQUE DES ACCENTS EN FRANCAIS

Bilal AL-SHAFI
Chercheur au Centre Tesnière
Faculté des Lettres de Besançon
Université de Franche-Comté

Le traitement informatique des accents et des signes auxiliaires dans un mot (1) peut se faire de deux manières :

1. En réalisant un dictionnaire électronique complet que nous pouvons consulter afin de vérifier l'orthographe d'un mot.
2. En trouvant des règles ou conditions simples qui gèrent l'utilisation des accents et des signes auxiliaires, puis en intégrant ces règles dans un programme afin de traiter les mots et les phrases à partir de nos connaissances sur le sujet .

Le premier procédé semble être simple car programmer un dictionnaire demande moins de travail que de trouver des règles. De plus, l'exhaustivité relative des dictionnaires assure de bons résultats à ce travail et une large utilisation. Cependant, ce travail sera limité par la définition même d'un mot (voir la définition ci-dessous). Un dictionnaire électronique ne peut pas traiter la conjugaison des verbes, et de ce fait les changements qui interviennent par exemple dans je cède, je céderai. Les possibilités de ce dictionnaire pour la correction automatique d'un texte sont très limitées ainsi que les possibilités d'utilisation didactique (apprendre par exemple l'orthographe ou la conjugaison...).

En revanche, chercher des règles, et ensuite construire un programme pour les exploiter, permet de comprendre le fonctionnement d'une langue, son évolution (dans le

(1) ("le mot est la plus petite unité qui correspond à un sens" ou "l'unité libre minimale" Grammaire Larousse du français contemporain page 11 : Jean- Claude Chevalier , Claire Blanche-Benveniste, Michel Arrivé , Jean Peytard, Paris (1988)