
Jabberwocky ou la symétrie dans les langues

Sylviane Cardey-Greenfield*¹

¹Centre Tesnière- Université de Bourgogne Franche-Comté – Université de Bourgogne Franche-Comté –
France

Résumé

Nous montrerons qu'il existe une certaine symétrie à l'intérieur d'une langue (intralangue) et que cette symétrie mise en évidence et modélisée va nous aider à traiter de façon automatisée les ambiguïtés de forme, les mots inconnus ou la néologie dans les textes.

Prenons un extrait de Jabberwocky de Lewis Carroll contenant des **mots** inventés par Lewis Carroll mais morphologiquement en adéquation avec la morphologie de la langue anglaise :

'Twas **brillig**, and the **slithy toves**

Did **gyre** and **gimble** in the **wabe**:

Un anglophone saura peut-être reconnaître que "brillig" et "slithy" sont des adjectifs même s'il ne comprend pas ce que ces mots signifient mais qu'en sera-t-il pour un ordinateur qui ne peut reconnaître que ce qui est stocké dans une base de données quelle qu'elle soit. De plus ici "toves" morphologiquement peut être nom ou verbe. Les néologismes sont un des points d'achoppement des systèmes de traitement automatique des langues, ils sont impossibles à déterminer par un ordinateur pour la bonne raison qu'ils ne sont pas dans les dictionnaires, ni répertoriés puisque ce sont des mots nouveaux.

Une recherche approfondie et systémique sur la morphologie des langues que nous avons menée nous a permis de mettre en évidence certains phénomènes de symétrie morpho-syntaxiques avec leur représentation formelle en intension basée sur la théorie des ensembles et la théorie des partitions. Grâce aux résultats de cette recherche que nous mettons en application de façon automatique sur des textes nous réussissons à reconnaître les parties du discours qu'ils contiennent (verbe, adjectif, nom, préposition, article...). Le plus intéressant est que nous pouvons par nos ensembles que nous faisons interagir les uns avec les autres, en respectant une partition du lexique et de sa morphologie que nous avons établie, déterminer la catégorie d'un mot inconnu, d'un néologisme, nous parlons dans ce cas de néologisme de forme.

Notre étude en morphologie ne se contente pas des règles de dérivation ou composition que l'on trouve généralement dans les manuels de grammaire ou de linguistique. Nous avons trouvé par exemple qu'en français tous les mots se terminant en -aal, -oal sont des noms. Nous avons mis au point tout un appareil formel, un modèle d'analyse qui traite les mots (nous parlerons de la notion de mot) d'une langue sous forme de règles organisées algorithmiquement. Une trace de toutes les opérations faites par la machine est accessible, ce qui en cas d'erreur nous permet d'apporter des corrections.

*Intervenant