

## ***‘Nouslisons’ : un logiciel référencier***

**BORDET Yves\***

**BOUHADIBA Farouk\*\***

**\*Centre Lucien Tesnière**

**Université de Franche-Comté,  
Besançon (France)**

**\*\* Université d’Oran (Algérie)**

### **Résumé**

Le logiciel *Nouslisons* donne l’âge auquel un texte est accessible pour un élève de langue maternelle et le niveau européen pour un étudiant de langue étrangère. *Nouslisons* existe pour le français, l’anglais, le russe et le chinois (mandarin). La méthodologie utilisée est la même pour toutes les langues et part d’un corpus littéraire. Cet article traite du logiciel *Nouslisons* pour l’espagnol et notamment du corpus d’étude.

### **1. Introduction**

Le logiciel *Nouslisons* permet de mesurer la difficulté d’un texte. Il donne l’âge auquel un texte est accessible pour un élève de langue maternelle et le niveau «européen» pour un étudiant de langue étrangère. Ce logiciel est établi selon une méthodologie bien précise. Il a été mis au point au Centre Tesnière de l’Université de Franche-Comté à Besançon. La méthodologie est celle utilisée au Centre Tesnière lors de l’établissement des logiciels *Nouslisons* pour le français, l’anglais, le russe et le chinois mandarin. L’hypothèse de départ est de considérer la langue littéraire comme étant la langue à étudier en priorité en évitant toute forme de langue spécialisée. Cette hypothèse tranche avec les différentes approches d’enseignement entreprises depuis les années 1950, mais renoue

avec l’approche traditionnelle de la pédagogie élaborée pendant l’Antiquité grecque et transmise jusque dans les années 1950. Les Grecs faisaient apprendre par cœur l’Iliade aux élèves de l’école primaire et abordaient ensuite l’étude de la grammaire puis de la rhétorique. Cette approche a été reprise intégralement par les Romains, puis transmise jusqu’à la Renaissance. Les mille ans du Moyen Âge ont vu un certain nombre de « renaissances » qui ont toujours été un retour à la tradition pédagogique grecque : renaissance bénédictine, irlandaise, carolingienne, clunysienne, scolastique, pour citer les plus marquantes, jusqu’à la grande Renaissance qui voit même la réapparition directe de l’étude des textes originaux de l’Antiquité. La période des Lumières perpétue la même tradition, et l’apparition de l’éducation gratuite, obligatoire pour tous (garçons et filles) fait de même. La première école gratuite, obligatoire pour tous (garçons et filles) apparaît dans la République de Genève en 1536. Elle est basée sur l’étude des textes classiques, et particulièrement de la Bible. L’école obligatoire française se met en place petit à petit et finalement lors des lois Ferry de 1881 et 1882. À l’époque, la grande majorité des Français ne parle pas français à la maison, mais patois ou langues régionales. Le français est appris et parlé à l’école. Les livrets de récitation ne contiennent que des textes littéraires, avec une majorité de textes en vers. Tous les textes de ces livrets de récitation, en vers et en prose, doivent être appris par cœur et récités devant le maître et devant la classe. On renoue ici avec la tradition grecque qui voulait que l’on développât le corps et l’esprit : pour les Grecs, « l’homme pense parce qu’il a une main », et « la langue est le muscle le plus puissant du corps ». Il faut donc maîtriser la main et la langue pour maîtriser tout le corps et l’esprit. On pourrait se poser la question : la disparition récente de l’apprentissage par cœur obligatoire de corpus littéraires et l’apparition récente de l’illettrisme sont-elles liées ? Et la réintroduction systématique de l’apprentissage

par cœur obligatoire de corpus littéraires permettrait-il de combattre l'illettrisme? Il s'agit de considérer (comme les Grecs) que l'étude de la grammaire doit être postérieure à l'acquisition de la langue. On doit apprendre la langue d'abord. Les capacités de curiosité et d'«absorbation» des élèves jeunes sont considérables. Si un élève de six à treize ans comprend 85 à 90% des mots d'un texte, il est capable de comprendre le texte assez rapidement après quelques répétitions et explications. La méthodologie d'établissement du logiciel *Nouslisons* part du traitement d'un corpus littéraire pour établir une base de données pouvant être traitée par l'informatique.

## 2. Méthodologie

### 2.1 Etablissement du corpus

Vingt textes en prose et vingt textes en vers choisis parmi les grands auteurs connus et reconnus dans la sphère où la langue est parlée et traduits dans des langues internationales officielles (français, anglais, espagnol, russe, arabe, chinois mandarin). Les textes font moins de 1000 mots. On essaie de prendre des textes en vers qui sont des poèmes et des fables complets. Les textes sont choisis parmi les auteurs les plus anciens jusqu'aux auteurs contemporains. Les textes sont mis en langue et orthographe modernes.

### 2.2 Exemple

Liste des 40 textes du corpus pour la langue française<sup>1</sup>.

---

<sup>1</sup>. BORDET Yves *Français littéraire et Français Fondamental, une étude lexicale* 2009 Thèse, Université de Franche-Comté, Besançon

XX TEXTES EN PROSE		
I	Michel Butor	La Modification
II	Louis- Ferdinand Céline	Voyage au bout de la nuit
III	Marcel Proust	Du côté de chez Swann
IV	Charles- Ferdinand Ramuz	La vie de Samuel Belet
V	Chateaubriand	Les Martyrs
VI	Gustave Flaubert	Madame Bovary
VII	Stendhal	Le Rouge et le Noir
VIII	Honoré de Balzac	Le Père Goriot
IX	Victor Hugo	Les Misérables
X	Georges Bernanos	Sous le soleil de Satan
XI	Voltaire	Pierre le Grand
XII	Jean- Jacques Rousseau	Les Confessions
XIII	Montesquieu	Les Lettres Persanes
XIV	Blaise Pascal	Les Provinciales
XV	La Bruyère	Les Caractères
XVI	René Descartes	Discours de la Méthode
XVII	Michel de Montaigne	Essais
XVIII	François Rabelais	Le tiers livre
XIX	Philippe de Commines	Portrait Moral de Louis XI
XX	Le Roman de Renart	La Pêche d'Ysengrin

XX TEXTES EN VERS		
I	Louis Aragon	Que serais-je sans toi
II	Philippe Jaccottet	Pensées sous les nuages
III	Paul Verlaine	Impression fausse
IV	Guillaume Apollinaire	Le Pont Mirabeau
V	Serge Gainsbourg	Oh! je voudrais...
VI	Jacques Prévert	Les feuilles mortes
VII	Paul Verlaine	Chanson d'automne
VIII	Arthur Rimbaud	Le dormeur du val
IX	Anonyme	À la claire fontaine
X	Victor Hugo	Demain dès l'aube

XI	Marceline Desbordes-Valmore	Les roses de Saadi
XII	Pierre de Ronsard	vous serez bien vieille...
XIII	Arthur Rimbaud	Ma bohème
XIV	Joachim du Bellay	D'un vanneur de blé aux vents
XV	Clément Marot	Dedans Paris
XVI	Charles d'Orléans	Le temps a laissé son manteau
XVII	José Maria de Hérédia	Les Conquérants
XVIII	Charles Baudelaire	Invitation au voyage
XIX	François Villon	Epitaphe Ballade des pendus
XX	Victor Hugo	Booz endormi

### 2.3 Remarque

Les textes et auteurs choisis ne sont pas des auteurs connus seulement régionalement ni par un phénomène de mode. Ils sont connus dans toute la francophonie et traduits dans d'autres langues. Ils sont reconnus pour leur valeur littéraire et enseignés dans les écoles officielles.

### 2.4 Traitement du corpus

Pour établir la liste de vocabulaire nécessaire à l'établissement de la base de données, on retient dans la liste un mot apparaissant au moins une fois dans au moins deux textes différents du corpus. La liste est appelée Français Fondamental Littéraire (en abrégé FFL).

### 2.5 Pour le français

Cette liste comporte 1502 mots. Chaque mot est ensuite décliné à toutes les formes qu'il peut prendre : masculin, féminin, singulier, pluriel pour les substantifs et adjectifs. Pour les verbes, toutes les formes simples conjuguées, y compris subjonctif imparfait et conditionnel passé. Par convention, on met la forme

infinitive en premier pour les verbes et la forme masculin singulier pour les noms et adjectifs.

## 2.6 Prise en compte de la ponctuation

Véritable respiration du texte, c'est un élément important et déterminant. Chaque signe de ponctuation est chargé d'une valeur particulière.

## 3. Résultats

### 3.1 Liste Français Fondamental Littéraire

Les études menées montrent que la compréhension d'un texte repose à plus de 75 % par la compréhension du vocabulaire<sup>2</sup>. Le logiciel *Nouslisons* donne le pourcentage de mots de la liste FFL dans un texte. L'analyse du corpus littéraire des 49 textes d'un manuel scolaire de 1909<sup>3</sup> donne des résultats étonnants. Ces 49 textes sont ceux du livret de récitation, c'est-à-dire les textes qui sont appris par cœur par les élèves pendant leur scolarité obligatoire de six à treize ans. Près de 90 % des mots de ces textes sont des mots de la liste du Français Fondamental Littéraire.

### 3.2 Vérifications

De nombreuses vérifications sont effectuées sur des corpus littéraires, techniques, scientifiques, articles de journaux, etc.<sup>4</sup>.

### 3.3 Le logiciel Lisi<sup>5</sup>

Le logiciel *Lisi* est un logiciel d'analyse de la lisibilité. Il donne le niveau scolaire des élèves qui peuvent comprendre un texte. La méthode d'établissement de la base de données est similaire

---

<sup>2</sup> MESNAGER Jean, BRES Stéphane *L'Atelier de lecture, Evaluer la difficulté des textes* 2008 Paris, Nathan

<sup>3</sup> BAUDRILLARD, BRELOT, BRISSET, DELPECH, et alii *Manuel pratique du Certificat d'Études primaires Livre de l'Élève* 1909 Paris, Delagrave.

<sup>4</sup> Bordet 2009

<sup>5</sup> Mesnager et Bres 2008

à celle de *Nouslisons* par certains côtés, mais diffère complètement par d'autres. Alors que *Nouslisons* procède à une approche résolument littéraire de la langue à enseigner, *Lisi* se base sur des listes de fréquence et des listes de vocabulaire enseigné selon l'année scolaire.

### 3.4 Comparaison des résultats

Comparaison de résultats de *Lisi* et *Nouslisons* pour les 49 textes du manuel 1909 : on a ici transposé en âge des élèves ce que *Lisi* traduit en niveaux scolaires français. Il est à remarquer que *Lisi* ne donne des résultats que pour des textes de plus de 300 mots. Lorsque ces textes ont moins de 300 mots, il donne seulement une indication de niveau. Nous avons marqué cette indication par un point d'interrogation.

	TITRE	Auteur	<i>lisi</i>	<i>Nouslisons</i>
1	<i>Égalité devant la mort</i>	Ronsard	11/12 ans?	11/12 ans
2	<i>Consolations à un père</i>	Malherbe	10/11 ans ?	10/11 ans
3	<i>Combat du Cid ...</i>	Corneille	9/10 ans?	10/11 ans
4	<i>Songe d'Athalie</i>	Racine	10/11 ans ?	11/12 ans
5	<i>Bataille de Rocroi</i>	Bossuet	11/12 ans	11/12 ans
6	<i>L'Avare</i>	Molière	10/11 ans	11/12 ans
7	<i>Les femmes savantes</i>	Molière	10/11 ans	10/11 ans
8	<i>La Mort et le Bûcheron</i>	La Fontaine	10/11 ans ?	10/11 ans
9	<i>Le Chêne et le Roseau</i>	La Fontaine	9/10 ans	9/10 ans
10	<i>Le Vieillard et les Trois Jeunes Hommes</i>	La Fontaine	10/11 ans	9/10 ans
11	<i>Le Loup et l'Agneau</i>	La Fontaine	9/10 ans	9/10 ans
12	<i>Le Coche et la Mouche</i>	La Fontaine	10/11 ans	9/10 ans
13	<i>Le Savetier et le Financier</i>	La Fontaine	8/9 ans	6 à 9 ans
14	<i>L'œil du Maître</i>	La Fontaine	10/11 ans	9/10 ans
15	<i>Les vœux</i>	Boileau	11/12 ans	11/12 ans
16	<i>Un grand seigneur</i>	De Sévigné	11/12 ans ?	9/10 ans
17	<i>Le madrigal de Louis XIV</i>	De Sévigné	10/11 ans	6 à 9 ans
18	<i>Diphile ou l'amateur d'oiseaux</i>	La Bruyère	11/12 ans	11/12 ans

19	<i>Portrait d'une jeune fille accomplie</i>	Fénelon	10/11 ans ?	10/11 ans
20	<i>La France en 1693</i>	Fénelon	10/11 ans ?	11/12 ans
21	<i>La fraternité</i>	Fénelon	10/11 ans ?	11/12 ans
22	<i>Un Persan à Paris</i>	Montesquieu	8/9 ans	9/10 ans
23	<i>La dot de la sœur</i>	Voltaire	10/11 ans ?	9/10 ans
24	<i>Sur l'état de nature</i>	Voltaire	10/11 ans	11/12 ans
25	<i>La vraie charité</i>	Rousseau	10/11 ans	10/11 ans
26	<i>Si j'étais riche</i>	Rousseau	13 ans et +	13 ans et +
27	<i>Amour filial</i>	Diderot	10/11 ans	6 à 9 ans
28	<i>L'âne</i>	Buffon	11/12 ans	10/11 ans
29	<i>L'Aveugle et le Paralytique</i>	Florian	8/9 ans	6 à 9 ans
30	<i>La mort choisissant un premier ministre</i>	Florian	10/11 ans ?	10/11 ans
31	<i>La carpe et les carpillons</i>	Florian	8/9 ans	6 à 9 ans
32	<i>L'amour du sol natal</i>	Chateaubriand	10/11 ans ?	13 ans et +
33	<i>La visite au toit paternel</i>	Chateaubriand	9/10 ans ?	12/13 ans
34	<i>Les soldats de l'an II</i>	Hugo	10/11 ans ?	11/12 ans
35	<i>Après la bataille</i>	Hugo	10/11 ans ?	9/10 ans
36	<i>Morts pour la patrie</i>	Hugo	8/9 ans	6 à 9 ans
37	<i>Le semeur</i>	Hugo	11/12 ans ?	12/13 ans
38	<i>Les victimes de la mer</i>	Hugo	9/10 ans ?	10/11 ans
39	<i>La mort du cheval</i>	Hugo	10/11 ans	10/11 ans
40	<i>La vie aux champs</i>	Lamartine	11/12 ans ?	13 ans et +
41	<i>Le pauvre colporteur</i>	Lamartine	10/11 ans ?	11/12 ans
42	<i>Retour</i>	Musset	10/11 ans ?	10/11 ans
43	<i>Le cor</i>	Vigny	10/11 ans ?	12/13 ans
44	<i>La fée</i>	George Sand	9/10 ans	9/10 ans
45	<i>L'alouette</i>	Michelet	13 ans et+ ?	13 ans et +
46	<i>Jeanne d'Arc</i>	Michelet	11/12 ans ?	10/11 ans
47	<i>Le travail</i>	Laprade	8/9 ans	9/10 ans
48	<i>L'amour de la patrie</i>	Renan	13 ans et + ?	13 ans et +
49	<i>Un troupeau de chèvres ...</i>	Taine	9/10 ans ?	11/12 ans

### 3.5 Remarques

Malgré deux approches totalement différentes, les deux logiciels donnent des résultats similaires. Il est évident que le faible

nombre de mots de la liste FFL est intéressant. Alain Bentolila<sup>6</sup> nous dit ceci:

A la fin du CE1 (âge 7 ans) les enfants au vocabulaire le plus pauvre ... connaissent une moyenne de 3000 mots radicaux (MR)<sup>7</sup>. ... le gain lexical annuel moyen après l'âge de 7 ans peut être estimé à 1000 MR par an ...

Le logiciel *Nouslisons* ne fait donc que confirmer que la méthode pédagogique héritée de l'Antiquité et transmise jusqu'à récemment est en fait la plus simple. Les méthodes basées sur les listes de fréquence et de vocabulaire ne font qu'appauvrir et compliquer l'enseignement.

#### 4. L'espagnol

##### 4.1 Le corpus

Le corpus espagnol a été choisi selon les mêmes critères que pour le français, le russe, l'anglais et le chinois mandarin. La situation est différente pour l'espagnol qui est parlé dans de nombreux pays du continent sud-américain et dont le système est différent de celui très centralisé de la France.

##### 4.2 La liste des 40 textes

### LES QUARANTE TEXTES

1	Don Juan Manuel	<i>El libro de Patronio o Conde Lucanor</i>
2	Fernando de Rojas	<i>La Celestina</i>
3	Anónimo	<i>El Lazarillo de Tormes</i>

<sup>6</sup> BENTOLILA Alain *Rapport de mission sur l'acquisition du vocabulaire à l'école élémentaire. A monsieur le Ministre de l'éducation nationale* 23 février 2007. Page 4.

<sup>7</sup> Mots radicaux : on ne prend pas en compte les mots obtenus par dérivation ou suffixation. (Note de Bentolila).

4	Miguel de Cervantes	<i>Don Quijote de la Mancha</i>
5	Mateo Alemán	<i>Vida del pícaro Guzmán de Alfarache</i>
6	Mariano José de Larra	<i>Una comedia moderna</i>
7	Fernán Caballero	<i>Los caballeros del pez</i>
8	Benito Pérez Galdos	<i>Fortunata y Jacinta</i>
9	Miguel de Unamuno	<i>El porvenir de España</i>
10	Vicente Blasco Ibañez	<i>La barraca</i>
11	Juan Ramón Jiménez	<i>Platero y yo</i>
12	Federico García Lorca	<i>La casa de Bernarda Alba</i>
13	Leopoldo Alas Clarín	<i>La Regenta</i>
14	Bio Baroja	<i>Zalacaín el Aventurero</i>
15	Juan Valera	<i>La primavera</i>
16	Emilia Pardo Bazán	<i>El árbol rosa</i>
17	Rubén Darío	<i>El rey burgués</i>
18	Fr. de Quevedo y Villegas	<i>El buscón</i>
19	Pedro Antonio de Alarcón	<i>Los ojos negros</i>
20	José María Pereda	<i>Los chicos de la calle</i>
21	Anónimo	<i>El Cantar del Mio Cid</i>
22	Juan Ruiz, Arcipreste de Hita	<i>El libro del Buen Amor</i>
23	El Marqués de Santillana	<i>Serranillas</i>
24	Jorge Manrique	<i>Coplas a la muerte de su padre</i>
25	Anónimo	<i>¿ Dónde vas Alfonso XII ?</i>

26	Pedro Calderón de la Barca	<i>Romance amoroso a una dama</i>
27	Garcilaso de la Vega	<i>Sonetos</i>
28	Luis de Góngora	<i>Terzetos</i>
29	Fray Luis de León	<i>A la vida retirada</i>
30	Lope de Vega	<i>¡ Oh libertad preciosa !</i>
31	Samaniego	<i>La cigarra y la hormiga</i>
32	José Zorrilla	<i>Don Juan Tenorio</i>
33	César Vallejo	<i>Los nueve monstruos</i>
34	Gabriela Mistral	<i>El encuentro</i>
35	Federico García Lorca	<i>Veleta</i>
36	Antonio Machado	<i>Cantares</i>
37	Gustavo Adolfo Becquer	<i>Volveran las oscuras golondrina</i>
38	Pablo Neruda	<i>Tango del viudo</i>
39	Miguel Hernández	<i>Aceituneros</i>
40	Rafael Alberti	<i>Sueño de marinero</i>

#### 4.3 Etablissement de la liste de l'Espagnol Fondamental Littéraire.

La même méthode a été utilisée. Est retenu un mot apparaissant au moins une fois dans au moins deux textes différents du corpus.

#### 4.4 Etablissement de la base de données

Comme pour le français, les variations des mots sont introduites avec toutes les formes verbales simples et toutes les variations possibles du mot. La programmation et la mise en place du

logiciel finalisé se fera en fonction des résultats d'études et de vérifications sur différents corpus littéraires, techniques, articles de journaux, etc.

Le prétraitement des unités lexicales du corpus sélectionné pour une analyse linguistique des textes en question se fait selon la même approche et la même procédure que pour le français (voir 1., Méthodologie), le russe, l'anglais, le chinois mandarin, l'espagnol, et l'arabe (en cours de réalisation).

Après la sélection de textes en prose et en vers, selon les mêmes critères que pour les langues cités *supra*, la première étape concerne l'étiquetage des unités lexicales et l'identification de leurs structures morphologiques, leur catégorie grammaticale et leurs fonctions syntaxiques et sémantiques. Ceci dans un souci d'éviter les ambiguïtés lexicales ou syntaxiques repérées dans le corpus et de localiser les anaphores et cataphores dans le corpus en place pour traitement et analyse. Cette phase se fait un peu plus tard dans le processus de traitement en élaborant des classes d'ambiguïtés et en utilisant l'approche micro systémique.

Pour ce qui concerne la langue arabe, nous essayons de constituer un corpus représentatif à partir d'une banque de données ouverte que nous alimentons au fur et à mesure. Les mêmes procédés et critères de sélection de textes en prose et en vers que pour les langues déjà mises en place (*voir supra*) ont été retenus pour le traitement de la langue arabe. Ceci, en tenant compte bien sûr des spécificités de cette langue. Nous citerons, à titre d'exemples le non-enchaînement des morphèmes de la langue arabe (du type non concaténatif), ses catégories grammaticales les plus productives, ses doublets sémantiques et surtout sa voyellisation / vocalisation qui est parfois source d'ambiguïté syntactico-sémantique et qu'il s'agit de repérer et d'inventorier.

Nous exposons ci-dessous un bref aperçu de étapes suivies et des résultats du prétraitement et de l'étiquetage pour l'espagnol et ce

après étude des textes espagnols en prose et en vers mentionnés en 4.2 ci-dessus.

#### 4.5 Prototypes de textes en espagnole traités automatiquement dans le cadre du logiciel «Nouslisons»

Texte I : (extrait de : *El libro de Patronio o Conde Lucanor*  
(Don Juan Manuel)

Lo° que° sucedió° a° una° zorra° con° un° cuervo° que° tenía° un° pedazo° de° queso° en° el° pico°. Hablando° otro° día° el° Conde+ Lucanor+ con° Patronio+, su° consejero\*, le° dijo°: Patronio+, un° hombre° que° se° llama° mi° amigo° comenzó° a° alabarme\* y° me° dio° a° entender° que° yo° tenía° mucho° poder° y° muy° buenas° cualidades\*. Después° de° tantos° halagos° me° propuso\* un° negocio\*, que° a° primera° vista° me° pareció° muy° provechoso\*

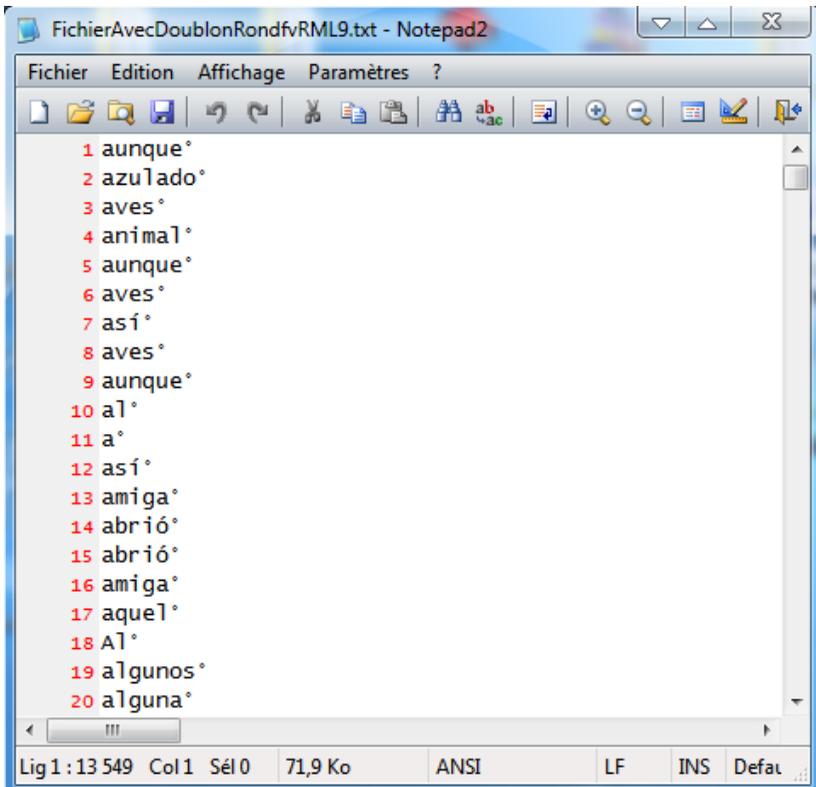
Poème I : (extrait de : *Sueño del marinero* (Rafael Alberti))

Yo°, marinero°, en° la° ribera° mía°,  
posada° sobre° un° cano° y° dulce° río°  
queda° su° brazo a° un° mar° de° Andalucía+,  
sueño° en° ser° almirante° de° navío°,  
para° partir° el° lomo° de° los° mar°es,  
al° sol° ardiente° y° a° la° luna° fría°.

De manière très succincte, la phase de prétraitement consiste essentiellement dans le repérage de mots utilisés deux fois et plus dans le texte intégral. Ces derniers sont étiquetés par un seul doublon rond (°) quelque soit leur catégorie grammaticale. Si le mot n'apparaît qu'une seule fois, il sera étiqueté par un astérisque (\*). S'il s'agit d'un nom propre (personne, lieu, etc.) qui apparaît deux fois et plus, il sera étiqueté par un plus (+). Une fois la fouille de corpus terminée (pour l'espagnol dans ce cas), nous procédons à la création de fichiers contenant des items

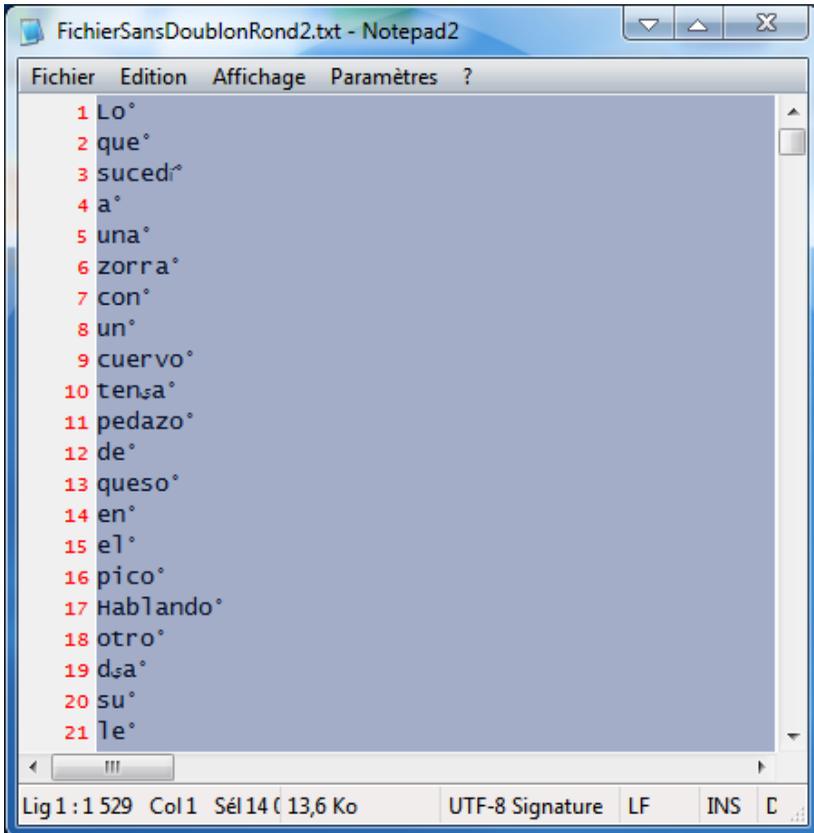
sans doublons ronds, des fichiers avec doublons ronds, etc. en faisant appel à des programmes informatiques pour le traitement des ressources et données obtenues.

### Exemple de Fichier avec doublon rond



```
Fichier Avec Doublon Rond fv RML9.txt - Notepad2
Fichier  Edition  Affichage  Paramètres  ?
1  aunque°
2  azulado°
3  aves°
4  animal°
5  aunque°
6  aves°
7  así°
8  aves°
9  aunque°
10 al°
11 a°
12 así°
13 amiga°
14 abrió°
15 abrió°
16 amiga°
17 aquel°
18 Al°
19 algunos°
20 alguna°
Lig 1 : 13 549 Col 1 Sél 0 71,9 Ko ANSI LF INS Defat
```

## Exemple de Fichier sans doublons rond



L'étape suivante consiste à prendre les données obtenues dans les fichiers sans doublons ronds et de la mettre sur Excel dans l'axe vertical. Les formes de bases sont identifiées pour chaque entrée lexicale. Par exemple, pour les verbes conjugués, la forme de l'infinitif est sélectionnée. Pour les noms et les adjectifs, c'est la forme masculin / féminin qui est retenue comme entrée lexicale de base. Les formes de base servent à recenser les variations grammaticales du mot en question.

Ainsi, pour un verbe à l'infinitif, toutes les formes conjuguées simples de la langue (l'espagnol dans ce cas) figurent sur l'axe horizontal (axe syntaxique). Autrement dit, les formes conjuguées à tous les temps simples du verbe en question et les formes du participe apparaissent sur Excel. Le même procédé s'applique aux noms et adjectifs avec leur forme de base suivie sur l'axe horizontal des formes masculin / féminin, masculin / singulier, féminin / singulier, masculin / pluriel, féminin / pluriel etc.

Nous présentons une illustration Excel pour le repérage des variables des formes de bases retenues. Le tableau ci-dessous représente un tableau beaucoup plus élaboré avec plusieurs entrées sur les axes vertical et horizontal. Il a été réduit au maximum pour pouvoir être inséré ci-dessous.

### Exemple de Tableau sur Excel

	A	B	C	D	E	F	G
1	Base Form	Verb	Masc.sg.	Fem.sg.	Masc.pl.	Fem. pl.	Adj.
2	de						
3	acostumbrado	acostumbrar	acostumbrad	acostumbrad	acostumbrad	acostumbrad	acostumbrado
4	a						
5	amigo		amigo	amiga	amigos	amigas	
6	aunque						
13	animal		animal		animales		
14	abrir	abrir	abierto	abierta	abiertos	abiertas	abierto
15	Así						
16	aquel		aquel	aquella	aquellos	aquellas	aquel
18	algunos		alguno	alguna	algunos	algunas	algunos
19	alguna		alguno	alguna	algunos	algunas	alguna
23	aquella		aquel	aquella	aquellos	aquellas	aquella
24	aliviar	aliviando	aliviado	aliviados	aliviada	aliviadas	alivio
25	alegres	alegrar	alegre	alegre	alegres	alegres	alegres
26	aquellos		aquel	aquella	aquellos	aquellas	aquellos
27	amores	enamorarse	enamorado	enamorada	enamorados	enamoradas	enamorado
28							

## Conclusion

Le logiciel *Nouslisons* démontre scientifiquement le bien-fondé de la pédagogie traditionnelle héritée des Grecs et abandonnée progressivement à partir des années 50 du vingtième siècle. Il est à remarquer que c'est à partir de cette époque qu'apparaît l'illettrisme, qui était inconnu auparavant. Les partisans de la pédagogie grecque ont été considérés longtemps comme retardataires et dépassés. L'approche résolument littéraire de l'enseignement dès l'école primaire est non seulement possible mais nécessaire et plus simple. *Nouslisons* en fait la démonstration scientifique. Les textes littéraires sont les plus simples et les plus accessibles, certains sont compréhensibles à des élèves de l'âge de six ans.

Nous espérons retrouver des résultats similaires pour les langues qui restent à analyser. Ceci surtout pour la langue arabe qui, comme cité supra, présente des caractéristiques spécifiques par rapports aux langues déjà analysées par le logiciel *Nouslisons*.

## Bibliographie

- BAUDRILLARD, BRELOT, BRISSET, DELPECH, et al  
*Manuel pratique du Certificat d'Études primaires Livre de l'Élève* 1909 Paris, Delagrave
- BENTOLILA Alain *Rapport de mission sur l'acquisition du vocabulaire à l'école élémentaire. A monsieur le Ministre de l'éducation nationale* 23 février 2007
- BORDET Yves *Français littéraire et Français Fondamental, une étude lexicale* 2009 Thèse, Université de Franche-Comté, Besançon
- BRUNO G. *Le Tour de la France par deux enfants* 1877 Paris, Belin (fac simile 1977 Paris, Belin)
- CHARTIER Anne-Marie (dir.) et HÉBRARD Jean (dir.)  
*Discours sur la lecture (1880-2000)* 2000 Paris, Fayard BPI Centre Pompidou

- CONSEIL DE L'EUROPE *Cadre européen commun de référence pour les langues* Conseil de l'Europe/Les éditions Didier, Paris 2006
- DELORS Jacques (président) *L'éducation, un trésor est caché dedans* Rapport à l'UNESCO de la commission internationale sur l'éducation pour le vingt et unième siècle, présidée par Jacques Delors 1996 Paris, Editions Odile Jacob
- DURKHEIM E. *L'évolution pédagogique en France* 1969 Paris PUF
- FURET François et OZOUF Jacques (Dir.) *Lire et écrire (T.1)* 1977 Paris, Les Editions de Minuit (avec le concours du Centre National de la Recherche Scientifique)
- FURET François et OZOUF Jacques (Dir.), Jeorger M., Nahoum V., Netter M. L., Pasquet Y., Compère M. M., Butel P., Mandon G., Poussou J.P., Lévêque P. *Lire et écrire (T.2)* 1977 Paris, Les Editions de Minuit (avec le concours du Centre National de la Recherche Scientifique)
- GAVOILLE Jacques, *Du maître d'école à l'instituteur*, 2010 Presses Universitaires de Franche-Comté
- GOUGENHEIM Georges, RIVENC Paul, MICHÉA R., SAUVAGEOT Aurélien *L'élaboration du Français Fondamental (1<sup>er</sup> degré)* 1964 Paris, Didier
- MESNAGER Jean, BRES Stéphane *L'Atelier de lecture, Evaluer la difficulté des textes* 2008 Paris, Nathan
- PALERMO J. *Histoire des institutions et des doctrines pédagogiques par les textes* 1958 Paris SUDEL
- PEYTARD J. (Dir.), BERTRAND D., BESSE H., BOURGAIN D., COSTE D., PAPO E., PELFRENE E., PORCHER L., SCTRICK R., *Littérature et classe de langue* 1982 Paris Hatier
- PLATON *La République* Œuvres complètes Bibliothèque de la Pleïade 1950 Paris, Gallimard

PONTY Janine *Polonais méconnus. Histoire des travailleurs immigrés en France dans l'entre-deux-guerres, Préface de J.-B. Duroselle* 1988 Paris, Publications de la Sorbonne  
SAINT AUGUSTIN *Les Confessions*

### Sites internet :

Coh-Metrix

Department of Psychology, University of Memphis, Memphis, TN-38152

<http://www.ncbi.nlm.nih.gov/pubmed/15354684>

Coh-Metrix calculates the coherence of texts on a wide range of measures. It replaces common readability formulas by applying the latest in computational linguistics and linking this to the latest research in psycholinguistics.

<http://www.scolarius.com/>

Scolarius est un outil gratuit d'analyse de la lisibilité. Il analyse le niveau de difficulté d'un texte en fonction de la longueur des mots, des phrases et des paragraphes.

*Analyse de la difficulté d'un texte français, la part du lexique*

MESNAGER J.

[http://www.google.ch/#hl=fr&scient=psy-](http://www.google.ch/#hl=fr&scient=psy-ab&q=analyse%20de%20la%20difficult%C3%A9%20d%27un%20texte%20fran%C3%A7ais&oq=analyse%20de%20la%20difficult%C3%A9%20d%27un%20texte%20fra&aq=0w&aqi=qw2&aql=&gs_l=hp.11.0.33i2112.1308.20551.0.22275.39.29.0.10.10.0.183.3300.0j29.29.0...0.0.z_GsuKhpGdQ&pbx=1&bav=on.2,or.r_gc.r_pw.r_qf.,cf.osb&fp=44f2c3614d6bbb2d&biw=1280&bih=687&pf=p&pdl=300)

[ab&q=analyse%20de%20la%20difficult%C3%A9%20d%27un%20texte%20fran%C3%A7ais&oq=analyse%20de%20la%20difficult%C3%A9%20d%27un%20texte%20fra&aq=0w&aqi=qw2&aql=&gs\\_l=hp.11.0.33i2112.1308.20551.0.22275.39.29.0.10.10.0.183.3300.0j29.29.0...0.0.z\\_GsuKhpGdQ&pbx=1&bav=on.2,or.r\\_gc.r\\_pw.r\\_qf.,cf.osb&fp=44f2c3614d6bbb2d&biw=1280&bih=687&pf=p&pdl=300](http://www.google.ch/#hl=fr&scient=psy-ab&q=analyse%20de%20la%20difficult%C3%A9%20d%27un%20texte%20fran%C3%A7ais&oq=analyse%20de%20la%20difficult%C3%A9%20d%27un%20texte%20fra&aq=0w&aqi=qw2&aql=&gs_l=hp.11.0.33i2112.1308.20551.0.22275.39.29.0.10.10.0.183.3300.0j29.29.0...0.0.z_GsuKhpGdQ&pbx=1&bav=on.2,or.r_gc.r_pw.r_qf.,cf.osb&fp=44f2c3614d6bbb2d&biw=1280&bih=687&pf=p&pdl=300)

*On retiendra de ce texte la présentation d'un outil d'aide à l'évaluation de la difficulté des textes, rapportée à la présence d'un vocabulaire plus ou moins connu des élèves.*

*SAINTE BASILE A des jeunes gens sur la lecture des Auteurs profanes* <http://www.abbaye-saint-benoit.ch/saints/basile/homelies/003.htm>

OFFICE QUÉBÉCOIS DE LA LANGUE FRANÇAISE

*Histoire de la ponctuation*

[http://66.46.185.79/bdl/gabarit\\_bdl.asp?id=3321](http://66.46.185.79/bdl/gabarit_bdl.asp?id=3321)