

Presenting the Nénufar project: A Diachronic Digital Edition of the *Petit Larousse illustré*

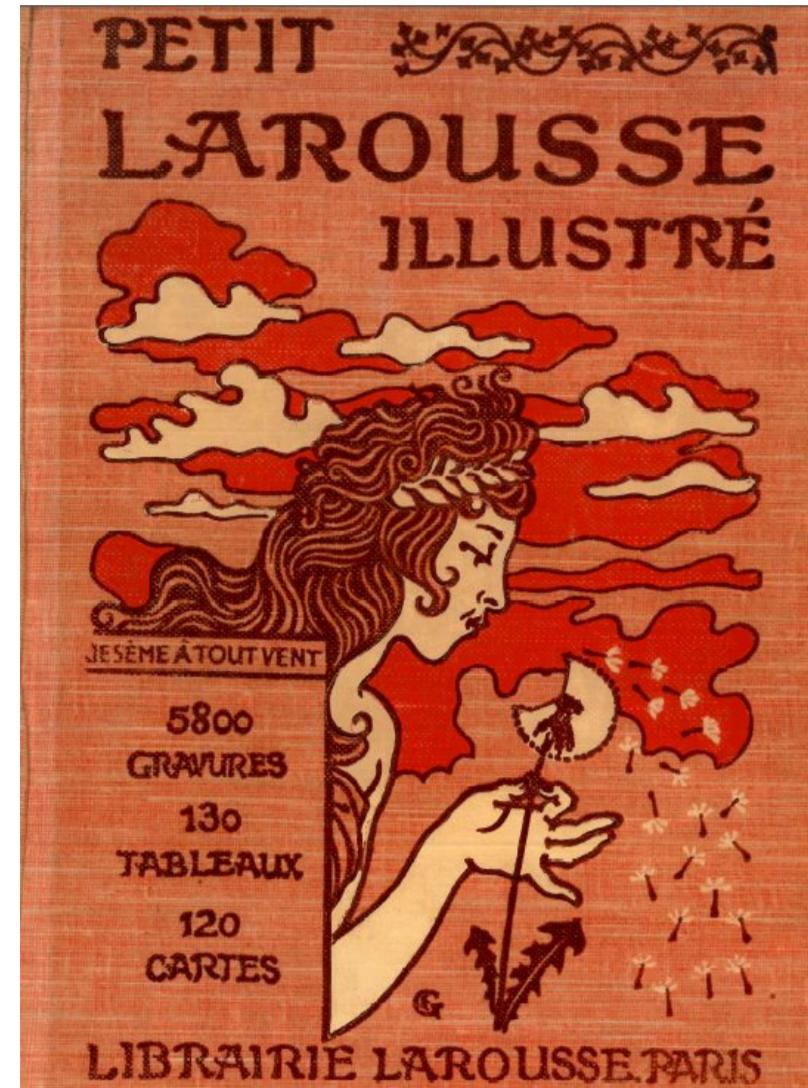


Hervé Bohbot, Alexandre Faucher, Francesca Frontini, Agata Jackiewicz, Giancarlo Luxardo, Agnès Steuckardt

Praxiling UMR 5267, CNRS – Univ. Paul-Valéry Montpellier 3, Montpellier, France

Mohamed Khemakhem, Laurent Romary

Inria – ALMAnaCH, Paris – Centre Marc Bloch, Berlin – Université Paris Diderot, Paris
Berlin-Brandenburgische Akademie der Wissenschaften, Berlin



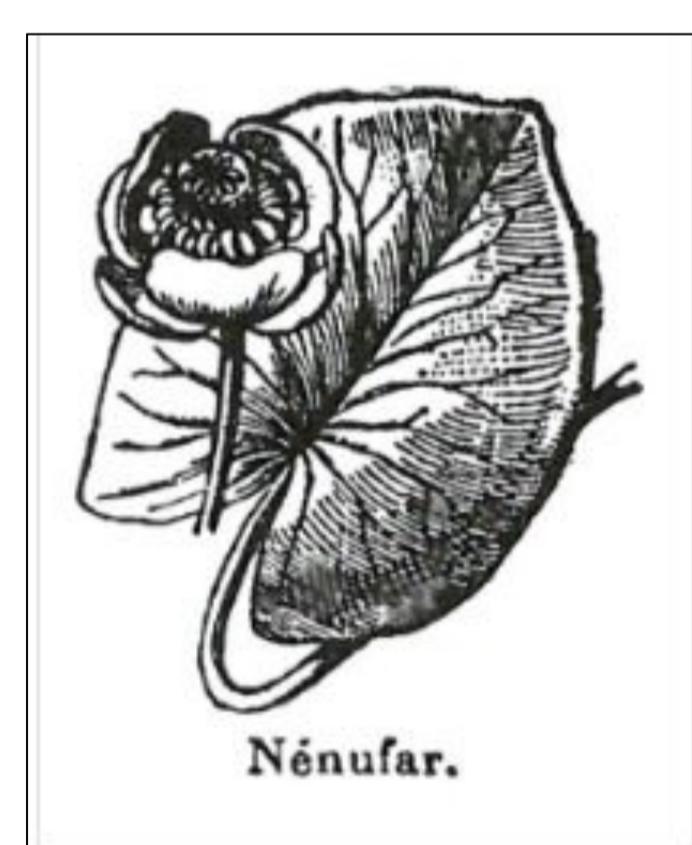
- The *Petit Larousse illustré* is the only French dictionary to be updated every year, since it was first published in 1905 (coined 1906).
- Handy and affordable, with each edition selling several hundreds of thousands copies (Pruvost, 2002), one *Petit Larousse* may virtually be found in every French household.
- Editions published prior to 1948 are public domain according to French laws, however, no open full text resource is currently available.

Language section (1906):

- 1066 pages *in-8°*
- 44876 entries
- 2532 illustrations
- ca 1.06 million words

Nénufar project: aiming to a diachronic corpus of 1906 to 1948 editions

Nénufar project name stands for *Nouvelle édition numérique de fac-similés de référence*. This name was chosen after a typically French polemic between supporters and opponents of some orthographic modifications of the French language. Amongst others, it was proposed that the waterlily (*Nymphaea L.*), should be spelled *nénufar* instead of *nénuphar*, which was perceived by some conservative people as an aggression towards the language integrity... However, it was actually the original, forgotten spelling, as it could be seen in not so old and common dictionaries like *Petit Larousse* (see below). There was clearly a gap to be filled, in order to bring to light recent evolutions of the language.



1906-1947 **NÉNUFAR** ou **NÉNUPHAR** n. m. Genre de nymphéacées aquatiques, à larges feuilles et à fleurs jaunes ou blanches qui croissent dans les tapis chauds et tempérés : le *nénufar blanc* est le lotus sacré des Egyptiens.

1948-1954 **NÉNUPHAR** ou **NÉNUFAR** n. m. Genre de nymphéacées aquatiques, à larges feuilles et à fleurs jaunes ou blanches, qui croissent dans les eaux douces et tempérées : le *nénuphar blanc* est le lotus sacré des Egyptiens.

NÉNUPHAR n. m. Genre de nymphéacées aquatiques, à larges feuilles et à fleurs jaunes ou blanches qui croissent dans les eaux douces et tempérées : le *nénuphar blanc* est le lotus sacré des Egyptiens.

From 1955

Some word datations:

- 1908: autobus, gréviste (*striker*)
1911: antimilitarisme
1912: puériculture, technicien
1914: tract
1915: copyright
1916: scientisme, syndicalisme
1918: boche (*pej. for German*)
1920: alerter
1921: bolchévik, bolchévisme
1922: mazout, pogrom, soviet

A new resource for studying recent history of language, culture and techniques

An example with the entry **AVIATION** between 1906 and 1912 editions

Nénufar

PRÉSENTATION GALERIE ABRÉVIATIONS SPARQL ENDPOINT

Rechercher, comparer, étudier les éditions de 1906 à 1924 du Petit Larousse Illustré.

⚠ Version de développement, données en cours d'ajout. Pour l'édition 1924, seule la lettre A est présente.

ex: hélicoptère...

Rechercher

Recherche avancée...

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z Pages roses A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Avezac-Macaya (Pascal d')
aviateur n. m.
aviateur, trice adj.
aviateur, trice n.
aviation n. f.
Avicenne
aviceptologie n. f.
avicle n. f.
aviculture n. f.
avide adj.
avidement adv.
avidité n. f.
Avienus
Avignon
Avila
Avila (Gilles-Gonzalès d')
Aviles
avillir v. a.
avillir (s') v. pr.

Édition - 1906-001

Article Historique Formes Occurrences Ressources

AVIATION (si-on) n. f. (du lat. avis, oiseau).

Vol des oiseaux. Navigation aérienne. — L'aviation désigne surtout la locomotion aérienne faite à l'aide d'un véhicule plus lourd que l'air. On a déjà fait de nombreuses tentatives à ce sujet, mais le problème n'est pas encore résolu. Les différents appareils qui ont été utilisés jusqu'ici peuvent être classés en trois catégories distinctes : les *orthoptères*, les *hélicoptères*, les *aéroplanes*.

Singulier aviation 40
Pluriel aviations 40
4
AVIATION n. f. machine aérienne. N. m. Machine aérienne. Signifie surtout l'usage d'un véhicule pour des oiseaux. N. signifie surtout l'usage d'un véhicule pour de nombreuses tentatives à ce sujet, mais le problème n'est pas encore résolu. Les différents appareils qui ont été utilisés jusqu'ici peuvent être classés en trois catégories distinctes : les *orthoptères*, les *hélicoptères*, les *aéroplanes*.

Rechercher les occurrences dans les articles aviation

2 articles contenant le mot "aviation":

AVIATEUR HÉLICOPTÈRE

Télécharger le TIEF

4 Afficher l'article au format XML

<entry xml:id='aviation' n='1906-001_1912-075'>
<form type='lemma'>
<orth>AVIATION</orth>
<pron>(<emph rend='italic'>si-on</emph>)</pron>
</form>
<gramGrp><pos>n.</pos> <gen>f.</gen></gramGrp>
<etym>(du <lang expand='latin'>lat.</lang> <mentioned>avis</mentioned>, <gloss>oiseau</gloss>).</etym>
<sense>
<def>Vol des oiseaux.</def>
</sense>
<sense>
<def>Navigation aérienne.</def>
<note type='encycl'>— L'aviation désigne surtout la locomotion aérienne faite à l'aide d'un véhicule plus lourd que l'air. On a déjà fait de nombreuses tentatives à ce sujet, mais le problème n'est pas encore résolu. Les différents appareils qui ont été utilisés jusqu'ici peuvent être classés en trois catégories distinctes : les *orthoptères*, les *hélicoptères*, les *aéroplanes*.</note>
</sense>
</entry>

References:

- Banski, P., Bowers, J., and Erjavec, T. (2017). TEI-Lex0 Guidelines for the Encoding of Dictionary Information on Written and Spoken Forms. In eLex 2017.
- Khemakhem, M., Foppiano, L., and Romary, L. (2017). Automatic Extraction of TEI Structures in Digitized Lexical Resources using Conditional Random Fields. In electronic lexicography, eLex 2017, Leiden, Netherlands, September.
- Bohbot, H., Frontini, F., Luxardo, G., Khemakhem, M., & Romary, L. (2018). Presenting the Nénufar Project: a Diachronic Digital Edition of the Petit Larousse Illustré. In GLOBALEX 2018 - Globalex workshop at LREC2018 (pp. 1–6). Miyazaki, Japan.
- Pruvost, J. (2002). *Les dictionnaires de langue française*. Coll. Que-sais-je ? n°3622, PUF Paris.

Nénufar project is headed by laboratoire Praxiling at Paul-Valéry University of Montpellier in collaboration with Inria. It is supported by the Délégation Générale à la Langue Française et aux Langues de France (an agency of French Culture and Communication Ministry) and the Huma-Num consortia CORLI and CAHIER.

Info: projet-nenufar@univ-montp3.fr

Data will be published under CC-BY license.