



**HAL**  
open science

## Base de données et histoire de la langue

Thierry Hatt

► **To cite this version:**

Thierry Hatt. Base de données et histoire de la langue. Bulletin de l'EPI (Enseignement Public et Informatique), 1987, 47, pp.168-174. edutice-00000990

**HAL Id: edutice-00000990**

**<https://edutice.hal.science/edutice-00000990>**

Submitted on 21 Oct 2005

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **BASE DE DONNÉES ET HISTOIRE DE LA LANGUE**

**Thierry HATT**

### **ÉTUDE DES CIVILISATIONS ET HISTOIRE DE LA LANGUE**

Le programme de classe de seconde des lycées d'enseignement général prévoit l'étude des "héritages" de la civilisation européenne. Il s'agit de montrer aux élèves qu'une civilisation est inscrite dans la longue durée, qu'elle puise à des sources très variées, que sa richesse est fondée sur sa diversité.

Les moyens pédagogiques utilisables pour montrer cette richesse et la diversité des influences sont nombreux : l'histoire des domaines politiques, religieux, techniques permet de décrire les stratifications et les choix qui fondent notre univers quotidien.

Pour sa part l'étude de la langue est un des moyens privilégiés de cette étude, pourquoi ?

La première raison est que tout enseignant est aussi professeur de français, qu'une telle étude est une occasion de l'enrichir, la deuxième est que les élèves sont particulièrement motivés par tout ce qui concerne les origines (la préhistoire par exemple les passionne toujours), ce dossier linguistique rencontre auprès d'eux un écho particulier, la troisième raison est que l'histoire de la langue permet aussi de traiter l'histoire générale. Les mots, en effet, sont venus avec les choses, dans des contextes d'influences que l'on peut décrire.

### **LA PRÉPARATION DU DOSSIER AVEC LES ÉLÈVES**

Au départ l'idée est simple : faire rechercher aux élèves dans les dictionnaires à leur disposition (étymologiques ou autres) une dizaine de mots "étrangers". On exclut d'emblée le vocabulaire latin et les termes scientifiques et techniques construits à partir du grec. A ma grande surprise les élèves reviennent dans les deux classes concernées avec des

listes de 50 à 100 mots par tête. Il faut donc s'organiser pour "traiter" un corpus aussi vaste. Nous décidons de recourir à l'informatique sur la suggestion d'une élève qui, avec l'aide de son père avait déjà créé une mini base avec son stock de vocabulaire.

On décide de faire figurer chaque mot sur une fiche séparée, d'y associer un thème, la date de première utilisation du mot, la langue de dernière origine.

Les doublons sont évidemment nombreux et lors de leur élimination en classe nous constatons que les dates fournies par les différents dictionnaires pour un même mot sont incompatibles et varient parfois de plusieurs siècles. En l'absence d'indications sur les critères qui ont présidé à leur choix, elles sont donc inutilisables. La recherche de dates cohérentes est reportée à une date ultérieure.

Les équipes de travail éliminent la plupart des doublons dans chaque classe mais il est impossible de les éliminer d'une classe à l'autre. Il est facile par contre de les éliminer par programme dans la base elle même.

### **Travail sur les thèmes**

Les thèmes associés aux mots sont un moyen de faciliter l'exploration de la base en limitant le domaine de recherche.

Cette étape du travail, menée collectivement, est essentielle. Il faut rendre cohérent le thésaurus des thèmes : par exemple utiliser uniquement des substantifs et pas d'adjectifs. Il faut hiérarchiser les thèmes utilisés de façon à placer dans la base de données un même niveau hiérarchique. Il faut par exemple choisir entre "économie" et "agriculture", "industrie", "services", entre "science" et une liste de sciences... Ce travail de sélection est intéressant avec des élèves à plus d'un titre. Il permet de préparer l'apprentissage du plan qui débouche en terminale sur la dissertation (Liste de thèmes jointe).

Le travail collectif détermine une trentaine de thèmes différents dont un thème "divers" qui posera des problèmes car 409 mots sur 1 600 vont s'y retrouver! Une prochaine étape de travail devra permettre de les reventiler.

## **L'étape informatique**

L'établissement dispose d'un bon système de gestion de base de données, jumeau d'un système très connu. Le professeur crée le format d'enregistrement ( mot, origine, thème) et les élèves au cours d'une séance de travail de deux heures saisissent leurs quelques dizaines de mots sur les huit postes disponibles.

Une fois la saisie terminée dans les deux classes il suffit d'ajouter toutes les saisies individuelles dans une base unique comme le permet le logiciel.

Le professeur élimine les doublons par programme, la base est alors réduite de 2 700 mots à 1 605. Deux index de tri sont créés, celui des thèmes et celui des origines ce qui accélère notablement les traitements ultérieurs. La base une fois corrigée et indexée est enfin recopiée sur huit disquettes différentes pour permettre les interrogations.

## **Exploitation de la base par les équipes de travail**

Plusieurs types de travaux sont possibles :

- Décomptes globaux :

Combien de langues sont représentées ? (58 en tout dont 17 bien représentées) Combien de mots en tout dans la base ? (1 605) Combien de thèmes ? (30) Le calcul de l'effectif des mots pour chaque langue et pour chaque thème donne lieu à la construction d'un histogramme .(figures 1, 2, 3 )

Ces tris simples sont réalisés soit avec les instructions disponibles dans le progiciel soit à la main sur les listings indexés. Les histogrammes sont d'abords dessinés à la main puis réalisés par programme sur imprimante. Pour améliorer la lisibilité les effectifs sont classés.

- un deuxième type d'interrogation consiste à s'intéresser soit à une langue particulière (l'arabe par exemple), soit à un thème relevant de plusieurs mots-clefs dans la base ( alimentation, cuisine, vie quotidienne).

Les listings correspondants sont édités sur imprimante avec des instructions du type LIST conditions multiples. Ces listings sont ensuite synthétisés sous la forme d'histogrammes. (figures 4, 5,6 )

Lorsque les équipes ont exploré tous les thèmes et toutes les langues, les résultats sont regroupés sous la forme de tableaux croisés. (thèmes X langues)

Quelles sont les conclusions de ce travail ?

Du point de vue des contenus, l'exploration de la base montre l'importance des apports étrangers : la richesse des apports arabes par exemple où plusieurs périodes peuvent être distinguées : scientifiques, techniques et commerciales du Moyen-Age (alidade, algorithme, alambic...), l'importance du relais espagnol (calebasse, laquais, massépin, abricot...), le rôle de la colonisation militaire au Maghreb (oued, toubib, barda...).

L'exploration de la base montre aussi les domaines d'élection de l'influence de chaque langue : à dominante "culturelle" pour l'italien, "militaire" pour l'allemand, de "vie quotidienne" pour l'anglais et l'américain.

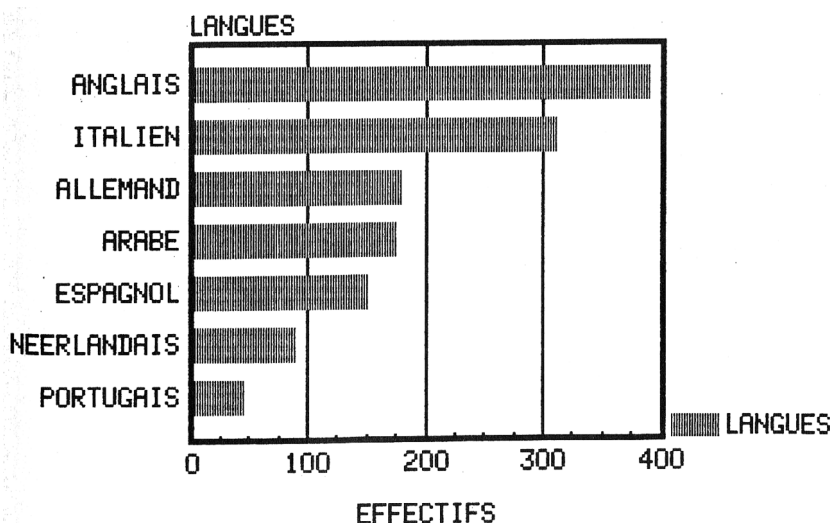
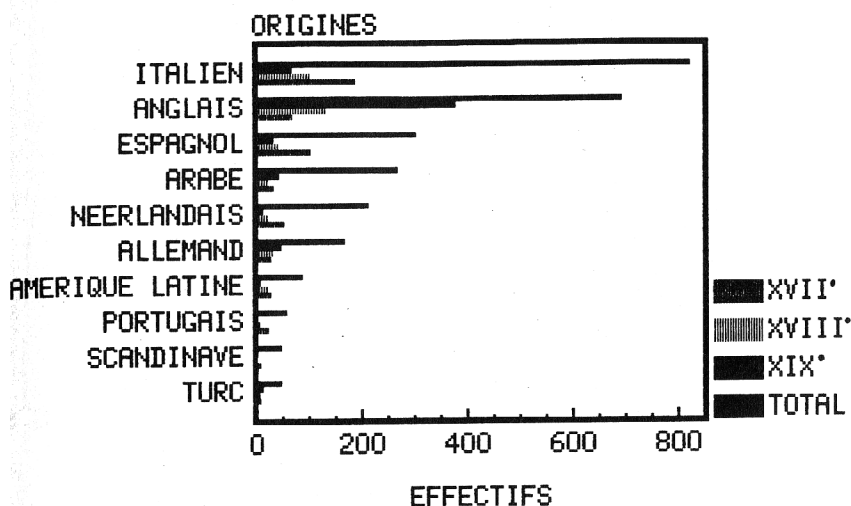
L'objectif pédagogique primordial de ce travail est la notion de PROJET. Il s'agit de faire réaliser à une classe toutes les étapes de recherche menant des données brutes aux résultats interprétés. Nous constatons chaque fois la grande efficacité pédagogique de cette démarche.

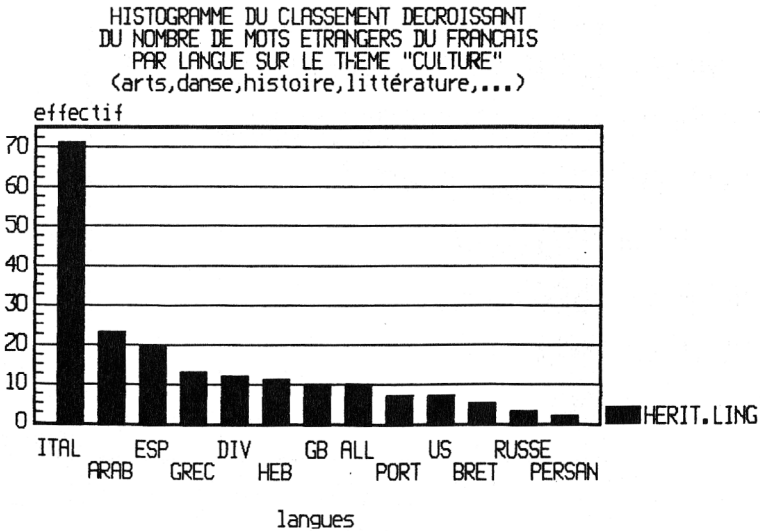
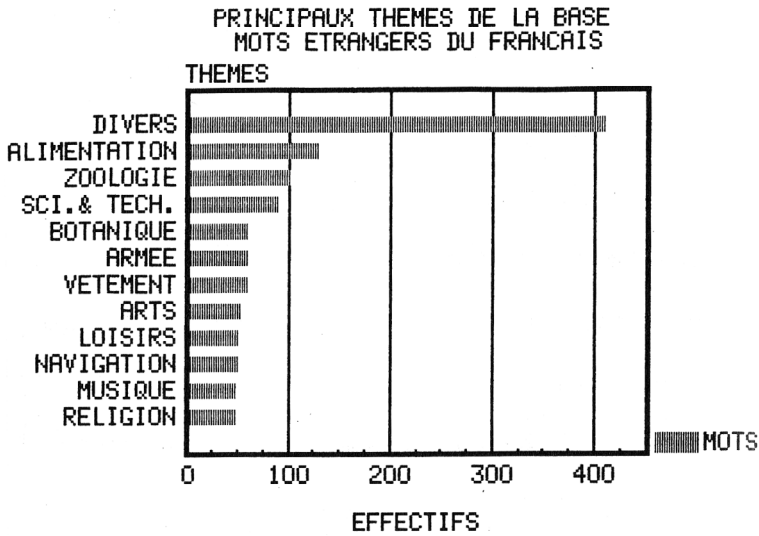
L'outil informatique est essentiel dans la mesure où la masse de données est telle qu'il n'est pas envisageable de la traiter à la main. L'ordinateur apporte aussi ses capacités graphiques et des résultats de qualité professionnelle.

Thierry HATT,  
Lycée Fustel de Coulanges Strasbourg,  
Juin 1987

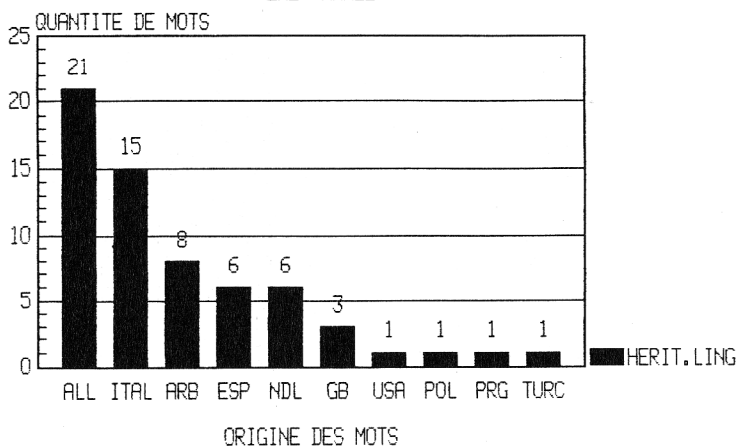
Référence : Pierre Guiraud, *Les mots étrangers*, QSJ n° 1166, PUF 1965.

NOMBRE DE MOTS "ETRANGERS" EN FRANCAIS  
D'APRES M. GUIRAUD  
( SUR UN TOTAL DE 2886 MOTS )

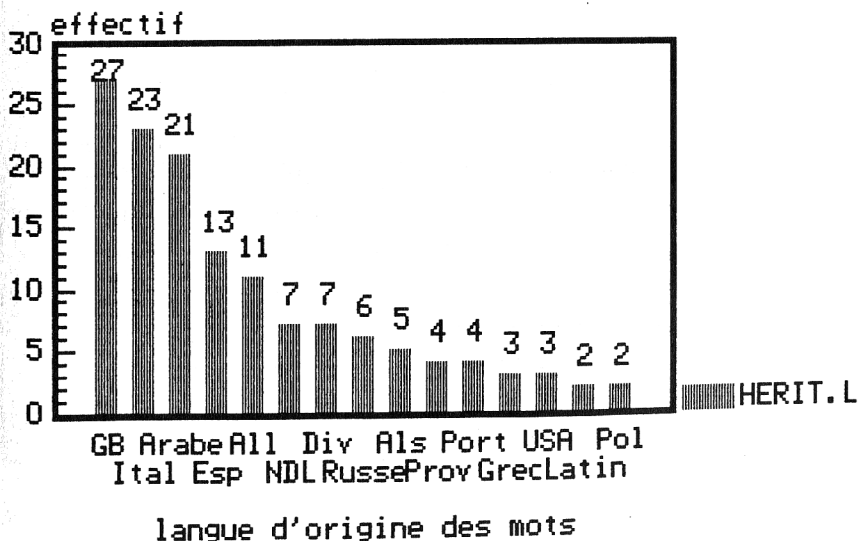




HISTOGRAMME DES MOTS FRANCAIS  
CLASSES PAR LANGUES  
THEME: ARMEE



Histogramme des effectifs des mots  
associés au thème "alimentation"  
mots clés: alimentation, cuisine, ...





them

ADMINISTRATI	LANGAGE
AGRICULTURE	LITTÉRATURE
ALIMENTATION	LOISIR
ANATOMIE	MATHÉMATIQUE
ARCHITECTURE	MÉCANIQUE
ARMÉE	MÉDECINE
ARMEMENT-OUT	MÉTALLURGIE
ARTS	MÉTÉOROLOGIE
ASTROLOGIE	MÉTIERS
ASTRONOMIE	MONDAIN
BOTANIQUE	MUSIQUE
CALENDRIER	NATURE
CIVILISATION	NAVIGATION
COMMERCE	OBJET
CRIME	OUTILS
CUISINE	PHILOSOPHIE
CULTURE	POÉSIE
DANSE	POLITIQUE
DIVERS	PROFESSION
DROIT	PSYCHOLOGIE
ÉCONOMIE	RELIGION
ÉNERGIE	SANTÉ
ETHNIE	SCIENCE ET T
FAUNE ET FLO	SOCIÉTÉ
FLORE	SPORT
GÉOGRAPHIE	TECHNIQUE
GÉOLOGIE	TISSUS
GUERRE	TITRES
HABITAT	TRANSPORT
HISTOIRE	VÊTEMENT
	VÊTEMENTS
	VIE QUOTIDIE
	ZOOLOGIE