



**HAL**  
open science

## Recueils de contes et nuages de mots.

Julien d'Huy

► **To cite this version:**

Julien d'Huy. Recueils de contes et nuages de mots.. Mythologie française, 2014, 255, pp.14-18.  
halshs-01099406

**HAL Id: halshs-01099406**

**<https://shs.hal.science/halshs-01099406v1>**

Submitted on 6 Jan 2015

**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.

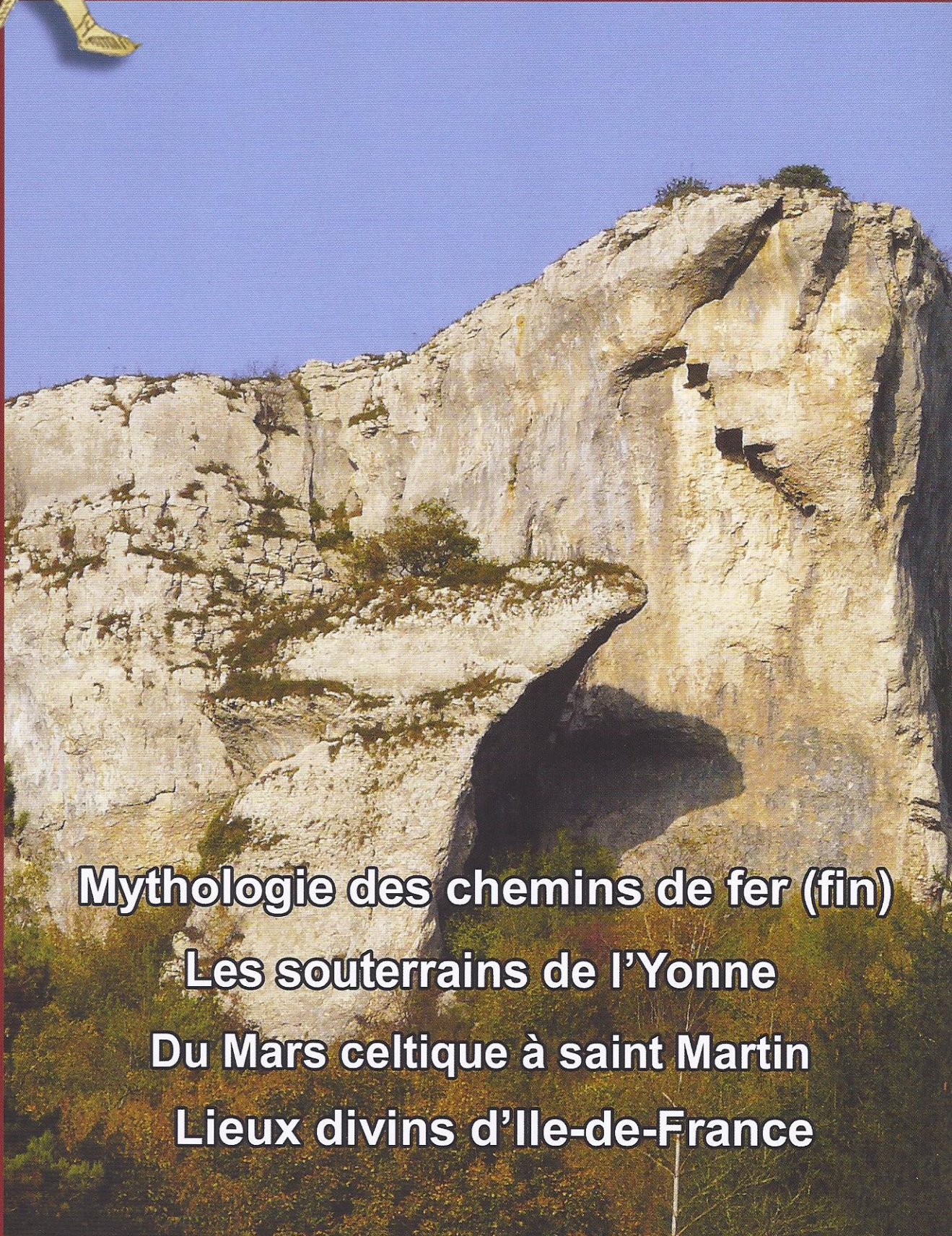


# Mythologie Française

Société de Mythologie Française

**Bulletin trimestriel n° 255**

**Juin 2014**



**Mythologie des chemins de fer (fin)**

**Les souterrains de l'Yonne**

**Du Mars celtique à saint Martin**

**Lieux divins d'Ile-de-France**

## RECUEILS DE CONTES ET NUAGES DE MOTS

par Julien d'Huy

Un nouvel outil heuristique a récemment été développé (Gambette et Véronis 2010) et utilisé pour l'analyse littéraire (Amstutz et Gambette 2010) : le nuage arboré. Il dispose les mots les plus fréquents d'un texte dans un arbre reflétant leur proximité sémantique calculée à partir de leurs cooccurrences. A notre connaissance, cet outil n'a jamais été employé pour étudier les corpus de mythes, alors que son pouvoir heuristique est certain. Afin de montrer comment ce type d'outils peut enrichir l'analyse en mythologie comparée, nous avons choisi de l'appliquer à quatre recueils de contes et légendes français.

Nous nous sommes appuyés sur les corpus occupant au moins deux volumes de la collection *Les Littératures populaires de toutes les nations*; autrement dit, les *Contes populaires de la Gascogne*, par Jean-François Bladé (1886), les *Contes populaires de Basse-Bretagne*, par François-Marie Luzel (1887), le *Folk-Lore de l'Île-et-Vilaine*, par Adolphe Orain (1897) et *Le Folk-Lore de la Beauce et du Perche*, par Félix Chapiseau (1902). S'appuyer sur un auteur et une collection par région permet une certaine normalisation du vocabulaire utilisé et, surtout, évite les redondances de récits. Nous avons utilisé les versions textes (*full text*) accessibles sur le site <https://archive.org>. Si la reconnaissance des lettres par le logiciel du site est parfois défectueuse, la masse de textes utilisée permet de contrebalancer ce défaut. Afin d'éviter les biais statistiques, nous avons par ailleurs supprimé les préfaces et, dans la mesure du possible, les titres de rappel en haut de chaque page.

Nous avons ensuite appliqué à ces données le logiciel en ligne *Treecloud* ([www.treecloud.org](http://www.treecloud.org)) créé par Philippe Gambette et Jean Véronis (2009). Nous nous sommes concentrés

sur les relations entretenues par les 75 mots les plus courants. Pour ce faire, nous avons adopté une formule de cooccurrence dite de Jaccard et avons programmé le logiciel afin que deux termes soient considérés comme liés si moins de cinquante mots les séparaient. Les groupements obtenus, dont nous discutons ci-après, ont été testés en changeant la formule de cooccurrence et le nombre maximal de mots, sans qu'aucune modifications significatives ne soient constatées. Plus la taille d'un mot est importante, plus souvent il est utilisé.

Les nuages de mots pour les quatre régions étudiées présentent une série d'oppositions : jour / nuit ; homme / femme (sauf pour la Basse-Bretagne) ; terre / eau (sauf pour la Beauce). D'autres oppositions semblent plus régionales : eau / feu, vieille / enfant, mère / fils, coucher / lendemain, nuit / matin ou matin / soir, etc. Cela tendrait à montrer que le conte, comme le mythe (Lévi-Strauss 1955), sert à rendre pensable des termes extrêmes grâce à la présence d'un terme moyen. Le fait que des termes antithétiques s'opposent régulièrement à quelques mots d'intervalle indique que ces oppositions structurent en profondeur le récit et sont perçues par les conteurs comme un ressort fondamental de l'action. Rappelons que la démonstration est ici statistique et que l'on ne peut faire l'économie d'une approche plus fine, chaque opposition étant en elle-même signifiante.

Le pôle essentiel, car se retrouvant développé dans nos quatre corpus, est celui de la famille<sup>(1)</sup> ; si le père peut en être absent, l'association mère-enfant s'y retrouve systématiquement (Ille-et-Vilaine : père, mère, fille, enfant, jeune ; Beauce / Perche : homme, femme, fille, enfant, mère, jeune, petit ; Basse-Bretagne : père, maison, grand, mère, fils, enfant, vieille, frère ; Gascogne : mère, grand, vieille, petite, petit, enfant). Une proximité sémantique entre la famille et la pauvreté (terme « pauvre ») existe également dans deux de nos corpus (Gascogne ;

1. Par pôle, nous entendons un nuage de mots localisés en un point spécifique de l'arbre et reliés par des relations de coprésence.

Basse-Bretagne), ce qui nous permet d'opposer le pôle familial et le pôle royal (ce dernier incluant des mots comme « palais », « cour » pour la Basse-Bretagne et « château » pour la Gascogne).

Le conte se construirait alors sur une opposition entre l'univers profane – celui de la famille et éventuellement de la pauvreté – et l'univers « autre », qui peut être royal (pôle important en Basse-Bretagne et Gascogne<sup>(2)</sup>) ou sacré (pôle important en Ille-et-Vilaine ; Beauce et Perche) ; on notera la présence du mot « porte » dans chacun des domaines « autres », ce qui nous permet d'interpréter cette dernière comme un lieu symbolique de passage. Pour la Beauce et le Perche, « porte » est ainsi associée à « église » et au mariage, donc à l'intégration des jeunes gens dans la communauté villageoise et religieuse. Tout se passe comme si le conte ramenait une multitude d'oppositions à une seule d'entre elles, sociale, palpable par les auditeurs et plus aisément dépassable.

Arrêtons-nous un instant sur le champ lexical du voyage. Dans les deux corpus où la royauté possède une place importante (Basse-Bretagne ; Gascogne), on remarquera que les mots liés au voyage (Basse-Bretagne : route, mit, partit, arriva, etc. ; Gascogne : partit, cheval, arriva, route, etc.) sont en connexion sémantique avec le domaine royal, impliquant que ce dernier ne se donne que comme aboutissement d'un chemin. L'insistance sur la prise physique du héros avec le monde se retrouve, dans nos deux corpus, avec la proximité des mots « tête » (qui pense) et « main » (qui agit).

Le roi est celui qui donne la main de sa fille (Ille-et-Vilaine : roi, amour ; Basse-Bretagne : roi, princesse ; Gascogne : roi, père) tout en étant lui-même en manque de quelque chose qui assure la continuité de la vie (Ille-et-Vilaine : roi, malade ; Basse-Bretagne : roi, vieux ; Gascogne : roi, vieux, mort) ; entre deux extrêmes (déficience / continuité), il est donc le point instable qui transmet le lien.

2. En Gascogne, le pôle royal se fond avec le champ lexical du temps qui passe : lendemain / matin / coucher / soir, etc.

La proximité entre le mot « femme » (ou « princesse ») et le mot « chambre » (Basse-Bretagne, Gascogne) et entre le mot « jeune » et le mot « fille » (dans les quatre corpus) peut s'expliquer par une volonté de lier la femme et la maternité, ou tout du moins la fécondité.

Pour nos deux corpus où l'ensemble de la royauté est important, on peut reconstruire le modèle suivant : un homme quitte sa famille (pauvre) et, à la fin d'un voyage, se voit donner la main d'une princesse par un roi, vieux ou malade, ce qui assure la continuité du lien social et permet d'apaiser les éventuelles tensions résultant d'une société de classes<sup>(3)</sup>. Ce modèle se trahit par une association suffisamment stable et récurrente (mais non systématique) de mots pour que l'algorithme les repère.

Notons par ailleurs que ces résultats font curieusement écho au témoignage de Jean-Louis Rolland, conteur breton étudié par Donatien Laurent ; à propos des contes, celui-ci expliquait qu'« il est difficile de les garder comme il faut dans la tête parce que beaucoup se ressemblent et qu'on tombe facilement de l'un dans l'autre. On y trouve les mêmes tours et les mêmes détours et pour finir ils s'enroulent en pelote et l'on n'arrive plus à les démêler. Or, c'est ça qu'il faut éviter parce que chacun suit sa propre route, même s'ils aboutissent tous au même point : le mariage du héros et de la fille du roi. » (Donatien 1981 : 37). Que beaucoup de contes « se ressemblent » indique une perméabilité des frontières entre histoires, et l'existence d'une trame commune à toutes, qui sous-tendrait au moins une grande partie des récits folkloriques bretons. « Tours » et « détours » renvoient au voyage. Jean-Louis Rolland conçoit d'ailleurs « chaque conte comme une sorte de voyage, (...) une sorte de circuit qu'il faut parcourir jusqu'au bout » (Donatien 1981 : 37). Enfin, comme nous l'avions supposé, ce voyage, quel qu'il soit, aboutit nécessairement au « mariage du héros et de la fille du

3. Comme le note Claude Seignolle à propos du folklore breton : « quand le fils du roi épousait une bergère, cela les vengeait du mépris de certains nobles ou de riches bourgeois à leur égard. » (1997 : 301).



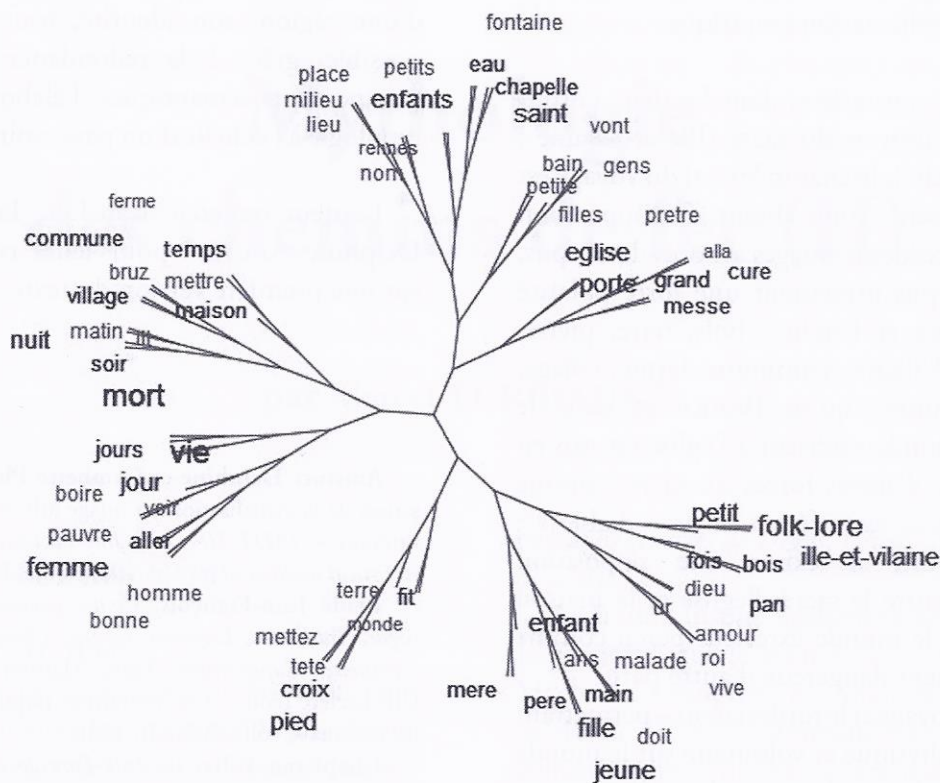


Figure 3. Graphe treecloud basé sur les ouvrages d'Adolphe Orain portant sur le folklore d'Ille-et-Vilaine.

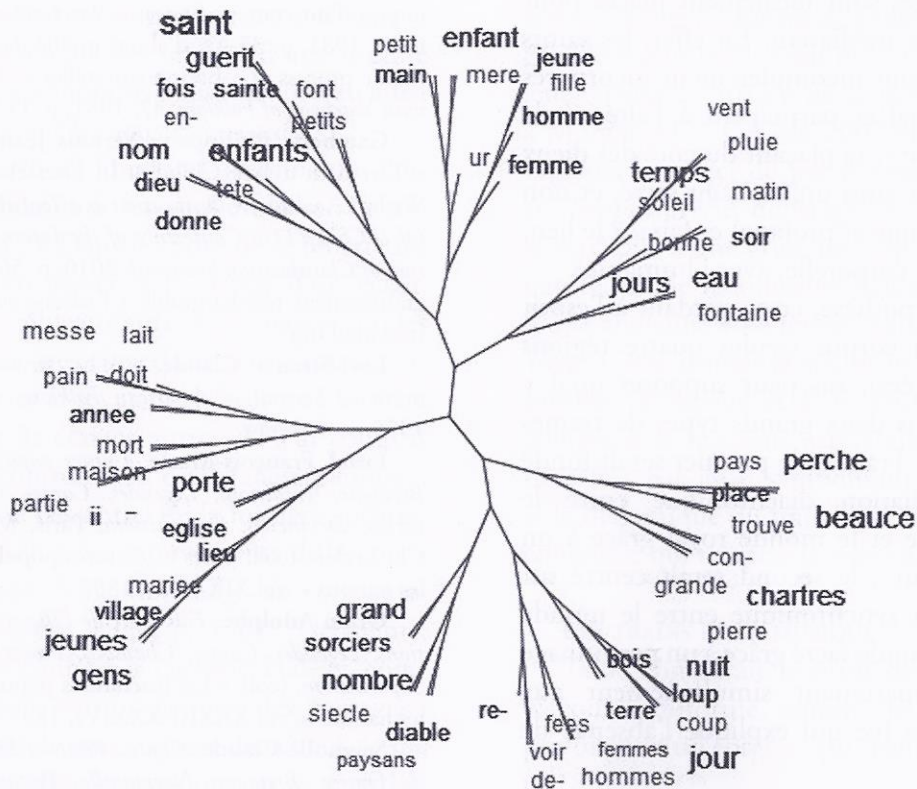


Figure 4. Graphe treecloud basé sur les ouvrages de Félix Chapiseau portant sur le folklore de la Beauce et du Perche.

roi ». Ici, l'approche statistique reçoit donc une étonnante confirmation empirique.

De façon contrastive, dans les deux corpus centrés sur l'univers du sacré (Ille-et-Vilaine ; Beauce / Perche), le champ lexical du voyage est très peu présent, voire absent. En opposition aux deux précédents nuages arborés du corpus, les deux corpus présentent une forte polarité locale (Beauce et Perche : bois, terre, pierre, etc. ; Ille-et-Vilaine : commune, ferme, village, maison). Notons qu'en Beauce et dans le Perche, ce monde extérieur à l'église est mis en relation avec d'autres forces du sacré, comme les « sorciers », les « fées » ou le « diable » ; autrement dit, il existe une opposition perceptible entre le sacré, l'église et la maison d'une part ; le monde extérieur perçu comme potentiellement dangereux d'autre part.

Là où le voyage et le roi donateur – permettant une action physique et volontaire sur le monde – servaient de médiateurs dans les corpus centrés sur la royauté, ce rôle semble attribué ici au « saint » et aux « enfants » (l'« enfant » au singulier étant connecté au noyau familial). Ces personnages sont idéalement placés pour jouer le rôle de médiateur. En effet, les saints et les enfants sont incomplètement incorporés au groupe social et participent à l'altérité de l'univers « autre », se plaçant du côté des dieux (la « porte » est ainsi un accès au sacré, et non à un lieu physique et profane) et faisant le lien, par leur nature corporelle, avec l'humanité.

A titre d'hypothèse, et en gardant à l'esprit la faiblesse du corpus (seules quatre régions ont été explorées), on peut supposer qu'il y aurait au moins deux grands types de trames folkloriques en France : le premier serait fondé sur une médiation diachronique entre le monde profane et le monde royal grâce à un voyage extérieur ; le second serait centré sur une médiation synchronique entre le monde profane et le monde sacré grâce à un personnage médiateur, appartenant simultanément aux deux catégories (ce qui explique l'absence de voyage).

L'utilisation de nuages de mots semble alors très prometteur. Une telle approche, pour peu qu'on la lie à une analyse plus empirique des

textes, permet d'approcher l'esprit folklorique d'une région, son identité, tout en rendant possible, grâce à la redondance de certains groupements sémantiques, l'élaboration d'une typologie à l'échelle d'un pays, voire du monde.

L'auteur remercie Jean-Loïc Le Quellec et Delphine Amstutz pour leurs commentaires sur une première version du texte.

\*  
\* \*

**Amstutz Delphine et Gambette Philippe**, « Utilisation de la visualisation en nuage arboré pour l'analyse littéraire », *JADT 10 : 10th International Conference on statistical analysis of textual data*, Rome, Italie, 2010.

**Bladé Jean-François**, *Contes populaires de la Gascogne, Traditions, Légendes, Contes, Chansons, Proverbes, Devinettes, Superstitions*, Paris, Maisonneuve Frères et Ch. Leclerc (coll. « Les littératures populaires de toutes les nations », vol.XIX-XXI), 1886.

**Chapiseau Félix**, *Le Folk-Lore de la Beauce et du Perche. Traditions, Légendes, Contes, Chansons, Proverbes, Devinettes, Superstitions*, Paris, J. Maisonneuve (coll. « Les littératures populaires de toutes les nations », vol. XLV), 1902.

**Laurent Donatien**, « Les procédés mnémotechniques d'un conteur breton », *Les Cahiers de Fontenay*, n°23, 1981, p. 35-42, d'abord publié dans « The mnemonic process of a Breton storyteller », *ARV, Scandinavian Yearbook of Folklore*, 37, 1981, p. 111-115

**Gambette Philippe et Veronis Jean**, « Visualising a Text with a Tree Cloud », In Locarek-Junge H. and Weihs C., editors, *Classification as a Tool of Research, Proc. Of IFCS'09(11th Conference of the International Federation of Classification Societies)*, 2010, p. 561-570, logiciel gratuitement téléchargeable à l'adresse suivante : [www.treecloud.org](http://www.treecloud.org)

**Lévi-Strauss Claude**, « The structural study of myth », *Journal of American Folklore*, vol.68, n°270, 1955, p. 428-444.

**Luzel François-Marie**, *Contes populaires de Basse-Bretagne, Traditions, Légendes, Contes, Chansons, Proverbes, Devinettes, Superstitions*, Paris, Maisonneuve et Ch. Leclerc (coll. « Les littératures populaires de toutes les nations », vol.XIX-XXI), 1887.

**Orain Adolphe**, *Folk-Lore de l'Ille-et-Vilaine, Traditions, Légendes, Contes, Chansons, Proverbes, Devinettes, Superstitions*, (coll. « Les littératures populaires de toutes les nations », vol.XXXIII-XXXIV), 1897.

**Seignolle Claude**, *Contes, Récits et Légendes des Pays de France. Bretagne, Normandie, Poitou – Charentes, Guyenne – Gascogne, Pays basque*, tome I, Paris, Omnibus, 1997.