

# LES SIGNES DE PONCTUATION COMME MARQUEURS DE SUBJECTIVITÉ DANS UN CORPUS LITTÉRAIRE

Vasilica MILEA-LE FLOCH  
Université Nancy 2

La présente analyse se propose d'étudier l'emploi des signes de ponctuation par Edgar Allan Poe dans ses nouvelles. Le repérage de la ponctuation permet également de réaliser une étude sur la longueur des mots et des phrases dans les nouvelles ainsi que sur leur degré de lisibilité. Ces opérations sont effectuées avec le logiciel *Hyperbase*, logiciel d'analyse statistique des données mis au point par le Professeur Étienne Brunet, de l'Université de Nice. Deux corpus sont soumis à l'analyse : *POE*, un corpus qui regroupe toutes les nouvelles d'Edgar Allan Poe et *EAF*, un corpus de référence constitué de textes de fiction américaine du XIX<sup>ème</sup> siècle.

La ponctuation est l'élément qui structure le texte et qui reflète la pensée du scripteur.

Si l'importance des signes de ponctuation n'est pas à démontrer, le passage suivant rappelle

leurs fonctions multiples dans le texte écrit :

When we speak English, we have all sorts of things we can use to make our meaning clear: stress, intonation, rhythm, pauses – even, if all else fails, repeating what we've said. When we write, however, we can't use any of these devices, and the work that they do in speech must be almost entirely handled by punctuation. (Robert L. Trask, *The Penguin Guide to Punctuation*, p. 2)

De plus, les signes de ponctuation font partie des traces de la subjectivité de l'énonciateur, car ils restituent à l'écrit les intentions de l'émetteur, son insistance ou le rythme qu'il veut imprimer à son discours. Ce ne sont que quelques exemples qui justifient l'intérêt accordé à ces marqueurs. Un troisième élément qui motive notre choix d'analyser les ponctèmes est l'intérêt que Poe lui-même manifestait pour ces outils :

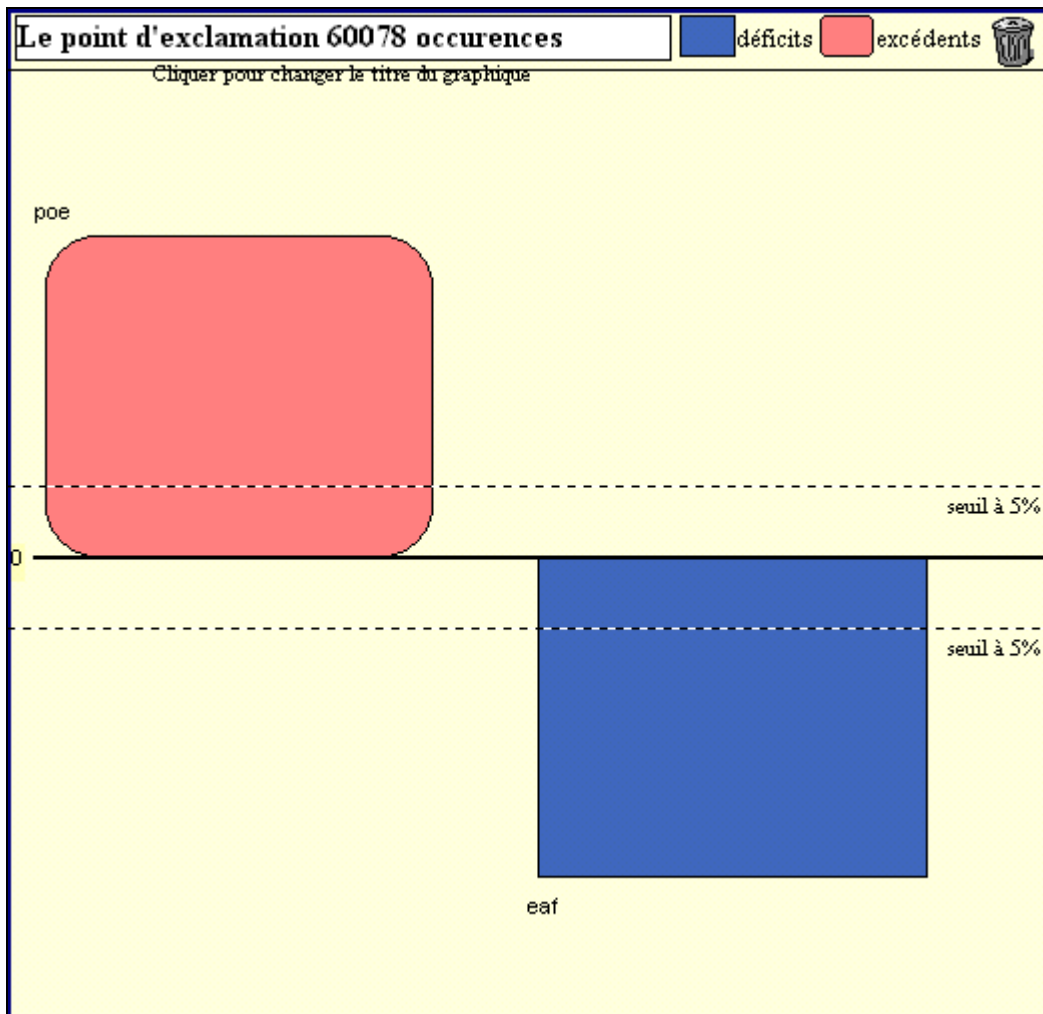
While we are upon the subject of style, we might as well say a word or two in regard to *punctuation*.[...] It cannot be said that the point is one of no importance - it is of very great importance. A slovenly punctuation will mar, in a greater or less degree, the brightest paragraph ever penned; and we are certain that those who have paid the most attention to this matter will not think us hypercritical in what we say. A too frequent use of the dash is the besetting sin of the volumes now before us. It is lugged in upon all occasions, and invariably introduced where it has no business whatever. (« Review of *Horse-Shoe Robinson* »)

La ponctuation semble avoir pour Poe deux fonctions distinctes : elle sert à ordonner sa pensée, mais aussi à décorer, en contribuant ainsi à l'iconicité textuelle. Le catalogue des signes de ponctuation ci-dessous permet de relever ce deuxième emploi, propre aux signes comme « \* » :

Corpus <i>POE</i>		Corpus de référence <i>EAF</i>	
	Ponctème Fréquence		Ponctème Fréquence
,	27523	,	1155185
.	12671	.	524007
-	9855	”	325391
“	5204	-	255392
;	2231	;	113462
!	1615	‘	112369
?	909	!	58463
‘	951	?	48235
(	531	:	16407
)	531	)	10007
:	289	(	10007
*	240	...	821
...	81	[	810
}	79	]	810
{	79	}	119
[	58	{	119
]	58	&	40
&	8	*	15
*†	2	/	13
/	4		

**Tableau 1.** Signes de ponctuation et leurs fréquences

La comparaison des deux corpus ne permet pas de relever de spécificités en ce qui concerne la fréquence. Il convient de s'intéresser au poids de chaque signe dans les corpus, afin de détecter d'éventuels emplois excédentaires. Le point, en tant que marqueur de la fin d'une phrase, présente un taux déficitaire dans *POE* par rapport à *EAF*. On observe la même situation pour le point d'interrogation. Ce déficit doit être mis en rapport avec la longueur des phrases. En ce qui concerne le point d'exclamation, le rapport est inversé, le corpus *POE* présente un taux excédentaire :



**Graphique 1.** Distribution du point d'exclamation

Parmi les signes de ponctuation utilisés à l'intérieur de la phrase, deux signes se remarquent par leur haute fréquence d'emploi. Il s'agit de la virgule et du tiret. Ce dernier est l'un des ponctèmes récurrents de l'écriture poésque ; sa fréquence théorique est de 5 257 occurrences, deux fois inférieure aux effectifs observés. Le tiret n'est pas dépourvu de signification :

Les énoncés entre parenthèses seraient comme des marques *d'altérité* au sein de sa propre voix, des marques de dénivèlement. C'est aussi la possibilité de ne pas explorer totalement un chemin, de ne pas tout éclairer de la même manière. (Francine Cicurel, « Le texte et ses ornements », p. 56.)

À la lumière de ces réflexions, le tiret et les parenthèses seront inclus et analysés en tant que signes porteurs d'une intention du scripteur, au même titre que le point d'exclamation, signe reconnu pour sa charge émotionnelle, donc subjective :

The exclamation mark (!), known informally as bang or shriek, is used at the end of a sentence or a short phrase which expresses very strong feeling. [...] exclamation marks are usually out of place in formal writing. (Robert L. Trask, *The Penguin Guide to Punctuation*, pp. 9-10)

Il convient de noter ici un autre aspect lié à la mise en évidence typographique et aux marquages pratiqués à l'intérieur des textes. Il s'agit de l'utilisation des majuscules et des italiques, techniques récurrentes chez Poe. Une première impression de lecture nous fait dire que les mots écrits en italiques sont très nombreux dans les nouvelles ; grâce au recours à *Hyperbase* on répertorie 431 formes et 1 128 occurrences. Le Tableau 2 ci-après permet d'établir la liste des mots marqués graphiquement par des italiques :

Fréquence Mot		Fréquence Mot	
87	the	9	to
27	a	9	never
25	of	7	be
19	in	6	no
16	not	6	my
15	was	6	all
14	very	5	you
13	must	5	now
12	it	5	me
11	I	5	did
		5	could

**Tableau 2.** Mots écrits en italiques dans le corpus POE

Tous les mots en italiques ayant une fréquence supérieure à 4 ont été inclus dans cette liste. De toute évidence, les mots grammaticaux ou les mots à fonction modalisatrice sont plus fréquents que les mots étrangers. La même constatation s'impose si on regarde la liste alphabétique des mots en italiques qui commencent par la lettre « a », présentée dans le Tableau 3 ci-après :

Fréquence	Mot	Fréquence	Mot
27	a	1	ammonia
1	à	1	am
2	abandon	1	amusement
1	about	15	and
1	ab	1	angel
1	absorbed	1	animal
1	accidents	1	animaux
1	acquainted	1	another
2	actually	3	an
4	acumen	1	any
1	added	1	apparent
1	Adolphe	1	April
2	after	7	are
1	age	1	arising
1	ah	1	ascertained
1	air	3	as
1	Alberto	2	Atlantic
1	Alexandre	3	at
1	Alfonzo	1	attention
1	Alger	1	attentive
1	alive	3	au
6	all	1	Austrian
1	ambois	1	autre

**Tableau 3.** Exemple de mots écrits en italiques dans le corpus POE

Sur un total de quarante-six entrées, seuls neuf termes et dix-neuf occurrences sont des mots d'origine étrangère ou des emprunts : *à, acumen, ambois, animaux, au, autre*, ou encore des noms propres : *Adolphe, Alberto, Alexandre, Alfonzo, Alger, April, Atlantic, Austrian*, ce qui représente seulement 17% des 112 occurrences.

Les italiques sont principalement employées pour la mise en relief ou l'emphase graphique d'un mot usuel. Les majuscules jouent le même rôle de démarcation et d'insistance : certains mots reçoivent un traitement de mise en valeur particulière, comme en témoigne leur orthographe. La liste des noms propres présents dans le corpus contient des noms comme ceux que l'on peut lire dans le Tableau 4 ci-après :

Mot	Fréquence	Mot	Fréquence
Madame	96	Pompey	36
Dupin	73	Goodfellow	35
Marie	72	Smith	35
Monsieur	63	Lalande	34
Jupiter	62	Roget	32
Legrand	47	Count	31
Paris	45	Etoile	30
Doctor	44	Lollipop	30
Sunday	43	Wyatt	30
East	40	April	29
Baron	37	John	29
Ellison	37	Easy	28
Grace	36	Ligeia	28
		Rotterdam	27

**Tableau 4.** Exemples de mots écrits avec majuscules dans le corpus POE

Figurent dans cette liste des mots qui ne sont pourtant pas des noms propres. Ce traitement particulier concerne surtout les noms communs qui acquièrent ainsi le statut de noms propres. L'auteur cherche à attirer l'attention du lecteur ; son intention semble évidente.

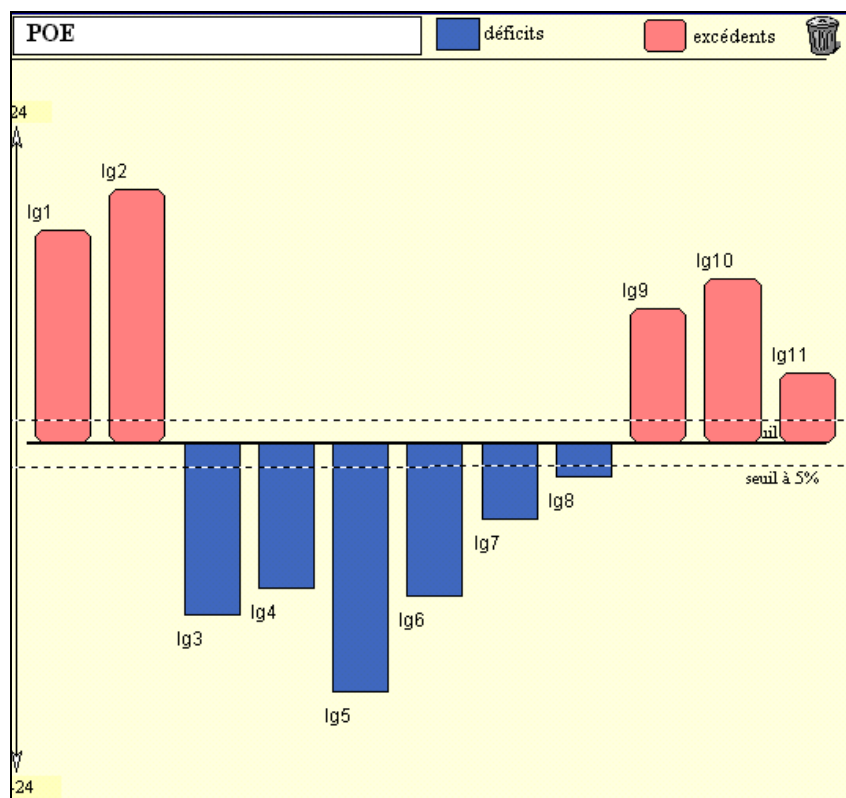
Il est difficile de définir statistiquement la phrase poesque, étant donné qu'elle peut varier en fonction du thème traité et du type de discours (direct ou indirect). Néanmoins, on peut calculer la longueur moyenne de la phrase, en divisant le nombre total d'occurrences du corpus par le nombre de signes de ponctuation de fin de phrase.

On obtient ainsi une longueur moyenne de la phrase poesque de 22 mots, valeur inférieure à celle constatée dans le corpus de référence EAF (24 mots par phrase). L'impression d'un style élaboré qu'on peut avoir à la lecture de nouvelles, ne se justifie donc pas par la longueur des phrases. Le mot en tant qu'unité minimale peut être analysé en fonction de sa longueur. *Hyperbase* permet de classer les mots d'un corpus en fonction de leur longueur en comptant leur nombre de lettres. Pour le corpus POE, la longueur des mots peut être présentée de façon synthétique de la manière suivante :

Longueur du mot	Occurrences
Longueur 1	14 940
Longueur 2	59 584
Longueur 3	68 783
Longueur 4	52 146
Longueur 5	32 564
Longueur 6	25 050
Longueur 7	21 176
Longueur 8	15 737
Longueur 9	19 242
Longueur 10	7 434
Longueur 11	609

**Tableau 5.** Longueur des mots dans le corpus POE

On remarque le taux élevé de mots courts, de une à trois lettres, et on sait que presque tous les mots grammaticaux entrent dans cette classe, ce qui justifie ces hautes fréquences. Pour les autres classes, la comparaison avec le corpus *EAF* est nécessaire, afin de détecter d'éventuelles spécificités. Le graphique suivant montre la distribution de chaque classe à l'intérieur du corpus, par rapport à *EAF* :



**Graphique 2.** Longueur des mots du corpus POE

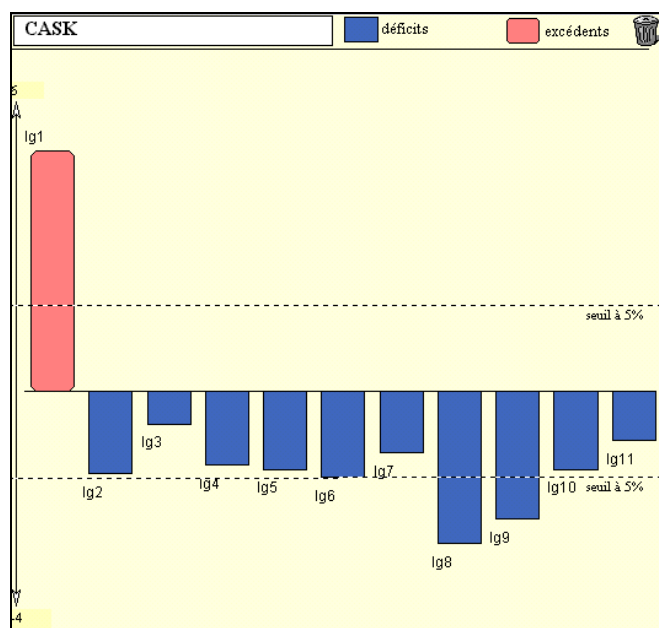
On constate ici que les mots longs sont excédentaires dans le corpus *POE*, signe d'une spécialisation linguistique accrue, comme le note Charles Muller :

Dans le cas du français moderne, le recours à un vocabulaire technique, savant, ou simplement recherché, augmente la proportion des mots longs [...] un style familier ou relâché sur le plan du lexique, agit évidemment en sens inverse. D'autre part, une syntaxe soignée, tend à réduire la densité du texte en mots grammaticaux, donc, en mots courts [...] (*Langue française et linguistique quantitative*, p. 153.)

L'école anglaise de la linguistique de corpus partage le même point de vue sur la présence des mots longs dans un corpus :

Conversations have frequent occurrences of present tense verbs, private verbs, first- and second-person pronouns, [...] coupled with relatively infrequent occurrences of nouns, prepositional phrases, long words [...] (Douglas Biber, *Dimensions of Register Variation: A Cross-linguistic Comparison*, p. 119.)

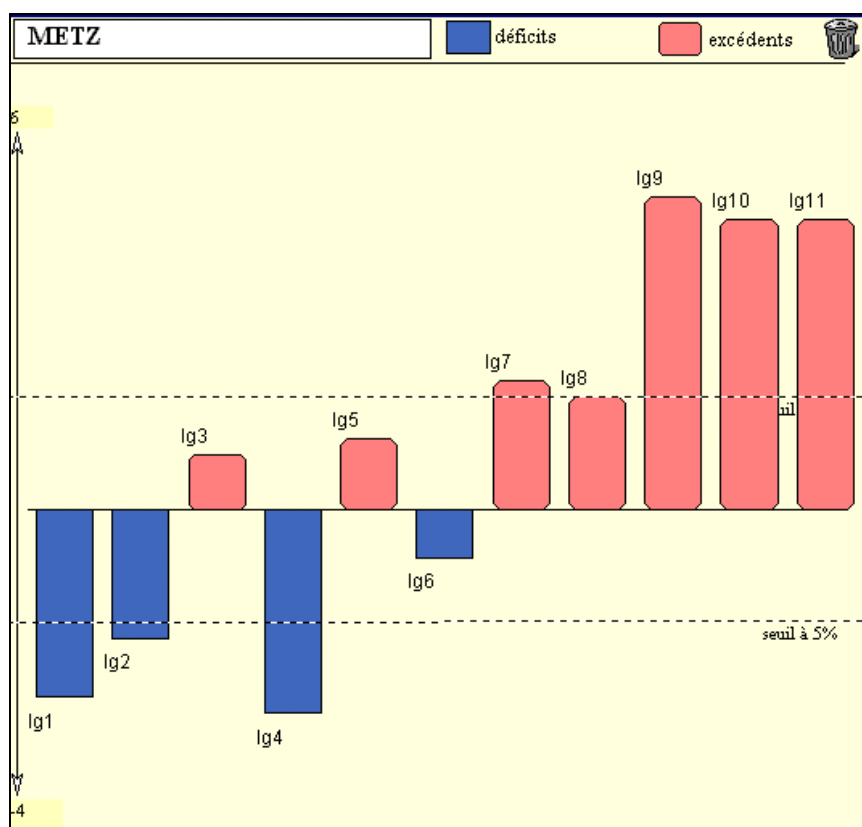
Biber donne donc la même explication pour l'anglais. Les mots longs sont l'apanage d'un style recherché à l'opposé du style familier et des conversations informelles. Le corpus *POE* met donc en œuvre un langage recherché, avec des mots longs et un taux élevé de noms et de substantifs. Cette opposition entre le dialogue et le style soutenu est visible dans les deux graphiques ci-après. Le Graphique 3 reflète la distribution des mots en fonction de leur longueur dans la nouvelle « The Cask of Amontillado » ; les déficits et les excédents sont calculés par rapport à l'ensemble du corpus :



**Graphique 3.** Longueur des mots dans « The Cask of Amontillado »



Le Graphique 4 ci-dessous présente les mêmes paramètres, mais pour la nouvelle « Metzengerstein », qui s'est déjà imposée pour sa richesse lexicale. Contrairement à « The Cask of Amontillado », qui est construite autour d'une situation dialogique et qui met en œuvre des mots courts, typiques de la langue parlée, la nouvelle « Metzengerstein » raconte les événements survenus dans la vie de deux familles hongroises, Metzengerstein et Berlifitzing. Le style est soutenu, les dialogues sont rares, et les mots longs, d'origine étrangère, abondent. Les différences de nature discursive entre les deux nouvelles se manifestent également au niveau des choix lexicaux, comme on peut le voir en comparant le Graphique 4 ci-dessous avec le Graphique 3 montré précédemment :



**Graphique 4.** Longueur des mots dans « Metzengerstein »

Le texte de la nouvelle permet d'observer que les passages de dialogue sont moins nombreux, mais surtout que les mots longs et peu courants abondent :

From this date a marked alteration took place in the outward demeanor of the dissolute young Baron Frederick Von Metzengerstein. Indeed, his behaviour disappointed every expectation, and proved little in accordance with the views of many a manœuvring mamma; while his habits and manner, still less than formerly, offered anything congenial with those of the

neighboring aristocracy. He was never to be seen beyond the limits of his own domain, and, in this wide and social world, was utterly companionless — unless, indeed, that unnatural, impetuous, and fiery-colored horse, which he henceforward continually bestrode, had any mysterious right to the title of his friend.

La longueur des mots ainsi que celle des phrases ont une influence directe sur la lisibilité du texte. De nombreuses études ont été consacrées aux facteurs de lisibilité, surtout du point de vue de la pédagogie et de l'enseignement.

D'autres facteurs affectent la lisibilité du texte ; il s'agit des italiques et des majuscules, mais ils ne sont pas pris en compte par ces tests. Il convient de garder à l'esprit que ces mots qui sont marqués graphiquement, sont porteurs d'une intention du scripteur et attirent l'attention sur eux-mêmes en tant que signes :

[...] parenthèses, énoncés entre tirets, usage de l'italique d'emphase, guillemets, etc. transmettent au lecteur une trace communicative émanant de l'auteur qui s'apparente à ces « ornements » musicales, dans le sens où ce marquage est l'indice d'un échange entre l'auteur (ou le compositeur) et le lecteur (ou l'interprète) et qu'il faut la coopération des deux instances pour qu'il y ait réussite de l'intention communicative ou esthétique. (Francine Cicurel, « Le texte et ses ornements », p. 53.)

Cette étude permet de démontrer que la subjectivité du texte écrit se manifeste non seulement au niveau verbal (choix lexicaux, modalisateurs, déictiques, pronoms personnels), mais également au niveau non verbal ou graphique, à travers tous les éléments qui constituent l'iconicité, le graphisme des textes.

En conclusion, nous pouvons affirmer que l'étude des signes de ponctuation et de l'emphase graphique s'inscrit tout naturellement dans une démarche qui cherche à identifier les manifestations de la subjectivité. Les ponctèmes constituent un élément structurant de la pensée poétique et témoignent d'une intentionnalité ; grâce à eux, le sujet signale sa présence dans le discours. Les exclamations sont des outils grâce auxquels le texte poétique exprime et transmet l'émotion et l'angoisse. Le point d'exclamation vise à produire un effet rhétorique et signale un apport subjectif. La lecture est ainsi rythmée, elle s'arrête et reprend au gré du scripteur qui parsème le texte de ces signes verticaux, repères visibles dans la linéarité de l'écriture. Un deuxième signe de ponctuation qui fonctionne comme un marqueur de subjectivité est le tiret. Ce ponctème intervient dans le texte pour l'éclaircir et

pour aider à sa compréhension, en introduisant une explication ou un ajout informationnel. C'est à ce niveau que la présence du sujet énonciateur se fait sentir, car ces ajouts sont autant de voies interprétatives proposées au lecteur. La lecture linéaire, automatique devient lecture active, exploration d'un texte qui héberge des indices textuels et graphiques.

## BIBLIOGRAPHIE

- Hyperbase, version 4.0.* Logiciel conçu par Étienne Brunet, Université de Nice, 1999.
- Biber, Douglas. *Dimensions of Register Variation: A Cross-linguistic Comparison.* Cambridge : Cambridge University Press, 1995.
- Brunet, Etienne. *Hyperbase : Manuel de référence, version 4.0.,* 1999.
- Cicurel, Francine. « Le texte et ses ornements ». *Figures d'ajout : Phrase, texte, écriture.* Authier-Revuz, Jacqueline. Lala, Christine. eds. Paris : Presses Sorbonne nouvelle, 2002.
- Muller, Charles. *Langue française et linguistique quantitative.* Genève, Slatkine, 1979.
- Poe, Edgar Allan. *Tales.* New York : Wiley and Putnam, 1845. <<http://www.eapoe.org/WORKS/editions/tales.htm>>. juin 2006.
- « Review of *Horse-Shoe Robinson* ». *The Southern Literary Messenger.* mai 1835, pp. 522-524. <<http://www.eapoe.org/WorkS/criticism/slm35k01.htm>>. juin 2005.
- Trask, Robert L. *The Penguin Guide to Punctuation.* London : Penguin Books, 1997.