

## Résumé Automatique Multilingue Expérimentations sur l'Anglais, l'Arabe et le Français

Houda Oufaida<sup>1</sup> Omar Nouali<sup>2</sup> Philippe Blache<sup>3</sup>

(1) Ecole Nationale Supérieure d'Informatique ESI, BP 68M Oued Smar, 16270, El Harrach Alger  
Algérie

(2) Centre de Recherche sur l'Information Scientifique et Technique CERIST, Rue Des 3 Frères Aissou,  
Ben Aknoun Alger Algérie

(3) LPL, AMU, CNRS, 5 avenue Pasteur, 13100 Aix-en-Provence  
h\_oufaida@esi.dz, onouali@mail.cerist.dz, blache@lpl-aix.fr

**Résumé.** La tâche du résumé multilingue vise à concevoir des systèmes de résumé très peu dépendants de la langue. L'approche par extraction est au cœur de ces systèmes, elle permet à l'aide de méthodes statistiques de sélectionner les phrases les plus pertinentes dans la limite de la taille du résumé. Dans cet article, nous proposons une approche de résumé multilingue, elle extrait les phrases dont les termes sont des plus discriminants. De plus, nous étudions l'impact des différents traitements linguistiques de base : le découpage en phrases, l'analyse lexicale, le filtrage des mots vides et la racinisation sur la couverture ainsi que la notation des phrases. Nous évaluons les performances de notre approche dans un contexte multilingue : l'anglais, l'arabe et le français en utilisant le jeu de données TAC MultiLing 2011.

**Abstract.** The task of multilingual summarization aims to design free-from language systems. Extractive methods are in the core of multilingual summarization systems. In this paper, we discuss the influence of various basic NLP tasks: sentence splitting, tokenization, stop words removal and stemming on sentence scoring and summaries' coverage. Hence, we propose a statistical method which extracts most relevant sentences on the basis of their terms discriminant power. We conduct several experimentations in a multilingual context: English, Arabic and French using the TAC MultiLing 2011 dataset.

**Mots-clés :** Résumé multilingue, analyse discriminante, TAL, évaluation multilingue.

**Keywords:** Multilingual summarization, Discriminant analysis, NLP, Multilingual evaluation.

### 1 Introduction

L'expansion du réseau internet à travers le monde a rendu disponible des documents écrits en différentes langues. En effet, le contenu web en anglais recule devant l'essor des autres langues ; Wikipédia.fr compte plus de un million et demi d'articles en mai 2014. Le contenu multilingue est donc en constante évolution et gagne de plus de plus de place sur le web. De ce fait, le développement d'outils multilingue est devenu indispensable : le résumé automatique multilingue en est un exemple.

Le but du résumé automatique est de produire une version condensée d'un ou plusieurs textes en utilisant des techniques informatiques. Ceci aidera le lecteur à décider si un document contient l'information désirée en un minimum de temps et d'effort. Récemment, de nouvelles tâches sont apparues et ont donné « un coup d'air frais » au domaine. Parmi ces tâches, on retrouve le résumé multilingue. Porté par les ateliers TAC MultiLing 2011<sup>1</sup> et ACL MultiLing 2013<sup>2</sup>, le résumé multilingue connaît une belle avancée grâce à la mise à disposition de corpus d'évaluation.

Cette tendance vient à l'encontre des premiers systèmes de résumé où l'idée était d'utiliser le traitement automatique des langues (TAL) dans le but de *reformuler* l'essentiel du texte source et de *générer* de nouvelles phrases:

<sup>1</sup> <http://www.nist.gov/tac/2011/Summarization/>

<sup>2</sup> <http://multiling.iit.demokritos.gr/pages/view/662/multiling-2013>

l'identification des paraphrases et la fusion d'informations en est un exemple (Barzilay, McKeown, 2005). Dans des cas plus rares, les techniques du TAL ont été utilisées pour produire des extraits tels que l'usage de l'analyse rhétorique RST (Rhetorical Structure Theory)(Marcu, 1998). Cependant, ces techniques requièrent des traitements de haut niveau souvent basés sur les ressources limitées et dépendantes à la langue.

Récemment, les approches statistiques ont prouvé leurs performances grâce à leur robustesse d'une part et à cause du manque d'outils TAL efficaces notamment dans un contexte multilingue d'une autre part. C'est dans ce cadre que s'inscrit notre proposition, notre système utilise un minimum de ressources linguistiques et est de ce fait facilement portable vers d'autres langues. Le processus d'extraction de phrases est purement statistique et repose sur les termes les plus discriminants. Nous utilisons ainsi une méthode d'analyse discriminante, mRMR (minimum Redundancy - Maximum Relevance) (Peng et al., 2005), pour la pondération des termes. Nous appliquons notre système sur trois langues : l'anglais, l'arabe et le français avec différents niveaux de traitements linguistiques: découpage en phrases, analyse lexicale, traitement des mots vides et racinisation.

Le reste de l'article est organisé comme suit : nous présentons dans la section 2 les principaux travaux dans le contexte multilingue notamment ceux des ateliers TAC Multiling 2011 et ACL Multiling 2013. Dans la section 3, nous décrivons notre méthode de pondération des termes et d'extraction de phrases. Nous discutons également les traitements linguistiques requis par notre système ainsi que leur niveau de dépendance à la langue. Dans la section 4, nous évaluons notre approche en utilisant un corpus multilingue et nous discutons les résultats obtenus. Enfin, nous terminons notre étude par une conclusion et quelques perspectives de recherche dans la section 5.

## 2 Travaux précédents

Le résumé multilingue consiste à concevoir des systèmes capables de résumer des documents écrits en différentes langues. Par conséquent, le système doit utiliser des techniques très peu dépendantes ou *idéalement* indépendantes de la langue source. A notre connaissance, (Mihalcea, Tarau, 2005) est le premier travail dans la thématique. Le système, TextRank, construit en premier lieu représentation graphique du texte où les nœuds représentent les phrases et les liens représentent la similarité entre ces phrases. Les auteurs définissent cette similarité, simplement, comme le nombre de tokens en commun. En deuxième lieu, les auteurs utilisent le principe de recommandation entre phrases en appliquant les deux algorithmes : PageRank et HITS. Ce travail a montré que l'utilisation d'un algorithme performant pour l'extraction de phrases avec un minimum de traitements liés à la langue peut conduire à des performances comparables à ceux utilisant des traitements linguistiques de haut niveau. Dans un contexte monolingue, (Ledeneva, 2008) affirme que les différents traitements linguistiques de base n'engendrent aucune amélioration des scores ROUGE (Lin, 2004) des résumés système en utilisant MFS (Maximal Frequent Sequences).

Dans un contexte multilingue, les tâches de base à savoir le découpage en phrases et la segmentation de ces phrases en un ensemble de tokens deviennent plus complexes. Jusqu'à présent, les chercheurs ont choisit d'appliquer diverses solutions, simplistes pour la plupart. (Litvak et al., 2010) utilisent une représentation vectorielle des documents et tentent de trouver la combinaison linéaire optimale entre 31 méthodes de notation de phrases. Les auteurs évaluent leur système sur deux langues : l'anglais et l'hébreu et utilisent un outil de découpage en phrases propre à chaque langue. (Boudin, Torres-Moreno, 2009) filtrent les mots vides et combinent entre la similarité cosinus et la mesure LCS (Longest Common Substring) basée sur les caractères pour la pondération des phrases. Les auteurs affirment que pour un seuil de similarité LCS supérieur à 0,6, le traitement remplace avantageusement la lemmatisation.

L'atelier TAC MultiLing 2011 évalue les difficultés liées à l'adaptation des méthodes au contexte multilingue. En effet, le corpus d'évaluation est un corpus parallèle de 7 langues. (Conroy et al., 2011) distinguent deux types de langues : simple casse et deux casses et utilisent un système de découpage en phrases adapté à chaque type: FASST-ONE et FASST-CAP. (Das, Srihari, 2011) utilisent le séparateur standard « . » tandis que (Hmida, Favre, 2011) utilisent le dernier caractère du document. (Saggion, 2011) utilisent les APIs GATE disponibles pour 4 langues. L'atelier ACL MultiLing 2013 reprend la tâche de l'atelier TAC MultiLing 2011, à savoir « Résumé multi-documents multilingue » et intègre trois nouvelles langues dont le Chinois. Pour l'analyse lexicale, (Conroy et al., 2013) distinguent 3 types de langues : anglais, non anglais et langues idéographiques (chinois), ils utilisent pour chacune une expression régulière particulière.

Le découpage en phrases et l'analyse lexicale sont des tâches de base du TAL et on peut soit utiliser des outils dédiés soit opter pour des solutions agressives et trop simplistes. Le problème avec le premier choix, est que ces outils ne sont disponibles que pour quelques langues (les plus utilisées). La tâche 2 de l'atelier ACL MultiLing 2013, résumé mono-document multilingue, vient tester la 2<sup>ème</sup> solution : simpliste. En effet, le corpus intègre un total de 40 langues et il

n'est évidemment pas possible de trouver des outils spécifiques à chaque langue. Pour cette tâche, (Conroy et al., 2013) classent les 40 langues en trois catégories et utilisent leur outil FASST-E adapté.

Les résultats de cette campagne, résumé mono-document multilingue, étaient plutôt faibles, aucun système n'a pu dépasser la Baseline pour l'anglais par exemple (Kubina, Conroy, & Schlesinger, 2013). Les organisateurs expliquent cette faible performance par le fait que les systèmes proposés jusqu'ici étaient un peu trop adaptés à la structure des articles de presse (pyramide inversée et l'utilisation du critère "position de la phrase").

Dans le présent article, nous proposons un système de résumé multilingue. Notre système utilise un minimum de traitements linguistiques et repose essentiellement sur des traitements numériques pour la pondération et l'extraction de phrases. En effet, le système procède en premier lieu au groupement des phrases en clusters. En deuxième lieu, les termes les plus discriminants vont être les mieux notés afin de mettre en avant les phrases saillantes. On utilise ainsi une méthode d'analyse discriminante: **mRMR** pour **Redondance minimale** et **Pertinence Maximale** (Ding, Peng, 2003). Nous proposons également, un nouvel algorithme d'extraction de phrases à deux vitesses : lente et rapide adaptable à la taille du résumé : court ou très court. La prochaine section détaille notre proposition.

### 3 Description du système

#### 3.1 Pondération des termes

Le but de l'analyse mRMR est la sélection d'un sous ensemble de variables qui représente au mieux l'espace total de variables. A chaque variable sont assignés deux scores: pertinence et redondance. Notre adaptation de cette méthode consiste à donner des scores à chaque terme de façon à sélectionner les termes les plus informatifs. Un terme est d'autant plus informatif si, étant donné les clusters de phrases similaires, sa fréquence varie significativement d'un cluster à un autre et au même temps n'attire pas un grand nombre de termes.

Le choix de l'algorithme de regroupement est primordial pour le succès ou l'échec de mRMR. En effet, la construction de groupes homogènes de phrases conduira naturellement à une meilleure estimation des scores. Ici, nous avons utilisé la classification ascendante hiérarchique et la similarité cosinus entre phrases. Bien entendu, l'utilisation d'un autre algorithme de classification/mesure de similarité est tout aussi possible.

On définit la pertinence d'un terme comme la valeur de l'information mutuelle entre ce terme  $t$  et la variable de classification  $h$  [1]. Si le terme et la variable de classification sont fortement corrélés alors le terme  $t$  est pertinent et nous permet ainsi de décrire au mieux les thèmes évoqués dans le texte.

$$Pertinence(t) = I(t, h) \quad (1)$$

Afin de calculer l'information mutuelle entre le terme et la variable de classification, il est nécessaire de construire la matrice des fréquences :  $M$  [Phrases  $\times$  Termes] où chaque ligne correspond à une phrase et chaque colonne représente un terme. La cellule  $M[i,j]$  contient la fréquence d'apparition du  $j^{\text{ème}}$  terme dans la  $i^{\text{ème}}$  phrase. A l'issue du groupement, chaque phrase devient rattachée à une classe bien spécifique. Ainsi, chaque ligne de la matrice est augmentée par la valeur de la variable de classification : le numéro de la classe.

Cependant, le fait qu'un terme décrit bien la variation des classes n'est pas suffisant. En effet, il faudra sélectionner des termes pertinents mais les plus dissimilaires possibles. On cherche ainsi à minimiser l'information mutuelle entre un terme donné et le reste des termes. La redondance d'un terme est définie par la moyenne de l'information mutuelle que ce terme partage avec les autres [2]:

$$Redondance(t) = \frac{1}{|T - \{t\}|} \sum_{j \in T - \{t\}} I(t, t_j) \quad (2)$$

On cherche ainsi à trier les termes de façon à maximiser la pertinence et minimiser la redondance. (Peng et al., 2005) proposent de maximiser soit la différence  $MID$  [3] ou le quotient  $MIQ$  [4]:

$$MID \stackrel{\text{def}}{=} \max_{t \in T} [Pertinence(t) - Redondance(t)] \quad (3)$$

$$MIQ \stackrel{\text{def}}{=} \max_{t \in T} [Pertinence(t) / Redondance(t)] \quad (4)$$

Le résultat de cette étape est le vecteur des termes avec leurs poids associés  $V_{mRMR} = (w_{t1}, w_{t2}, \dots, \dots, w_{tn})$

### 3.2 Pondération des phrases

Etant donné le vecteur  $V_{mRMR}$ , la tâche est d'attribuer des scores à chaque phrase selon les poids mRMR de ses termes. Pour ce faire, nous calculons la moyenne des poids mRMR des termes de la phrase [5]:

$$Score(P) = \frac{1}{n_p} \sum_{i=1, n_p} w_{t_i}, t_i \in V_p \quad (5)$$

### 3.3 Extraction des phrases

Dans notre système, le terme est l'unité de calcul des scores et la phrase est l'unité d'extraction. Le résumé sera ainsi formé de  $n$  phrases dans la limite de la taille du résumé requise (nombre de caractères, nombre de mots ou pourcentage de compression). Nous définissons deux vitesses d'extraction : rapide et lente. La première, rapide [6], remet à zéro les poids des termes dans  $V_{mRMR}$  dès leur intégration au résumé.

$$\forall t \in \{V_{mRMR} \cap V_p\}, w'_t = 0 \quad (6)$$

La deuxième, lente [7], diminue le poids des termes selon le score de la dernière phrase choisie ( $P$ ). La première vitesse convient dans le cas où on doit générer des résumés très courts car sinon on risque de sélectionner du bruit. La deuxième stratégie sélectionne progressivement les phrases et donne plus de *temps de vie* au terme.

$$\forall t \in \{V_{mRMR} \cap V_p\}, w'_t = w_t - sim(V_{mRMR}, V_p) * w_t \quad (7)$$

### 3.4 Contexte Multilingue

Dans le contexte multilingue, la dépendance de la langue doit être minimale. En effet, dans le système que nous proposons la dépendance est située au niveau du prétraitement. Notre approche requiert au *minimum* un découpage *correct* en phrases et une analyse lexicale et au *maximum* le filtrage des mots vides et la racinisation. Dans nos expérimentations, nous allons étudier l'impact de l'intégration de ces tâches de manière plus ou moins poussée.

## 4 Evaluation

Le corpus TAC MultiLing 2011 (Giannakopoulos et al., 2011) contient 10 groupes de documents à résumer. Chaque groupe contient à son tour 10 articles de presse décrivant une séquence de nouvelles autour du même événement: les attaques à la bombe de Londres en 2005 ou le tsunami de 2004, etc. Les textes dans les autres langues ont été traduits de l'anglais par des locuteurs natifs en suivant l'approche phrase par phrase. Trois résumés modèles sont fournis dont la taille est comprise entre 240 et 250 mots.

### 4.1 Protocole d'évaluation

Notre objectif étant l'évaluation de l'impact de la qualité des traitements linguistiques appliqués sur notre méthode: découpage en phrases naïf ou adapté à la langue source, segmentation en mots ou utilisation d'un tokenizer, le filtrage des mots vides et enfin l'application ou non de la racinisation. Nous avons pu tester notre approche sur trois langues : l'anglais, l'arabe et le français dans la limite de la disponibilité des outils. Le tableau suivant décrit les différents outils utilisés pour chaque langue :

	Découpage en Phrases	Analyse lexicale	Filtrage des mots vides	Racinisation
Anglais	Stanford tokenizer	Stanford tokenizer	Liste de 571 mots	Porter Stemmer
Arabe	Expression régulière	Expression régulière	Liste de 168 mots	Shereen Khodja Stemmer
Français	Europarl LinguaSentence	Expression régulière	Liste de 127 mots	Snowball Stemmer

TABLE 1 : Outils TAL pour l'anglais, l'arabe et le français

Nous avons donc défini 4 niveaux d'analyse, et pour chaque niveau nous avons généré les résumés en utilisant notre approche:

- **SR**: Découpage en phrases naïf (arabe) ou adapté (anglais, français)
- **TSR**: SR+Analyse lexicale adaptée (anglais, français) ou simple segmentation en mots (arabe)
- **WSR**: TSR+Filtrage des mots vides
- **SSR**: WSR+ racinisation

## 4.2 Résultats et Discussion

Les tableaux suivants montrent les résultats ROUGE-1 et ROUGE-2 (Lin, 2004) des résumés produits par notre système. Il est à noter que les résumés ont été générés avec un groupement à 8 clusters, vitesse rapide.

	ROUGE-1	ROUGE-2		ROUGE-1	ROUGE-2		ROUGE-1	ROUGE-2
	F1-score	F1-score		F1-score	F1-score		F1-score	F1-score
<b>CLASSY</b>	<b>0,48361</b>	<b>0,17612</b>	<b>CLASSY</b>	<b>0,34296</b>	<b>0,14207</b>	<b>JRC</b>	<b>0,43539</b>	<b>0,17199</b>
<b>BASELINE</b>	<b>0,40343</b>	<b>0,11445</b>	<b>BASELINE</b>	<b>0,26788</b>	<b>0,08982</b>	<b>BASELINE</b>	<b>0,37324</b>	<b>0,09346</b>
<b>SR.MRSYN8</b>	0,42056	0,11647	<b>SR.MRSYN8</b>	0,2743	0,09492	<b>SR.MRSYN8</b>	0,37324	0,10428
<b>TSR.MRSYN8</b>	0,43531	<b>0,12739</b>	<b>TSR.MRSYN8</b>	<b>0,2893</b>	0,09736	<b>TSR.MRSYN8</b>	0,39482	0,11074
<b>WSR.MRSYN8</b>	<b>0,44111</b>	0,12512	<b>WSR.MRSYN8</b>	0,28125	<b>0,10044</b>	<b>WSR.MRSYN8</b>	<b>0,39508</b>	<b>0,11917</b>
<b>SSR.MRSYN8</b>	0,4398	0,12671	<b>SSR.MRSYN8</b>	0,27196	0,08494	<b>SSR.MRSYN8</b>	0,3463	0,0935
<b>Anglais</b>			<b>Arabe</b>			<b>Français</b>		

TABLE 2 : Résultats ROUGE-1 et ROUGE-2 pour l'anglais, l'arabe et le français

On constate que les meilleurs résultats sont obtenus avec le filtrage des mots vides. Ils dépassent la Baseline mais restent en dessous du meilleur système. De plus, la différence entre les 4 variantes n'est pas significative. Nous avons donc évalué le nombre de phrases en commun entre les résumés R1 et R2 issus de deux niveaux consécutifs à l'aide de la formule suivante [8]:

$$Accord(R1, R2) = \frac{2|R1 \cap R2|}{|R1| + |R2|} \quad (8)$$

Le tableau suivant présente les résultats de l'estimation de la moyenne d'accord entre les résumés produits à chaque niveau d'analyse avec le niveau qui le suit:

	Arabe	Anglais	Français	Moyenne
<b>SR×TSR</b>	47,09%	39,78%	39,74%	42,20%
<b>TSR×WSR</b>	38,86%	32,98%	40,91%	37,58%
<b>WSR×SSR</b>	30,78%	44,94%	50,69%	42,13%

TABLE 2 : Variation des phrases extraites avec les niveaux de traitements TAL

On constate qu'avec un seul traitement à la fois, en moyenne 40% des phrases sont maintenues. La plus grande différence (ou désaccord) est enregistrée pour la langue arabe avec l'application du racinisateur. Le plus grand accord est enregistré pour le français avec l'application du racinisateur Snowball. Pour l'anglais, la plus grande différence est enregistrée avec le filtrage des mots vides. Ceci est dû à la taille de la liste des mots vides particulièrement longue pour cette langue. Ainsi, malgré les performances ROUGE très proches des résumés issus de chaque étape (SR, TSR, WSR et SSR), le niveau d'accord est en moyenne 40% et ne dépasse pas un maximum de 50%.

Il est également intéressant de noter que les différents traitements effectués ont un impact non négligeable sur la notation et par conséquent le choix des phrases à extraire. Plus de 70% des phrases choisies au départ (SR) ne le sont pas après traitement (SSR). On remarque aussi qu'avec la racinisation le constat est à l'opposé pour l'arabe et pour le français : on y enregistre la plus grande différence pour l'arabe et le plus grand accord pour le français. En analysant les textes après racinisation, on remarque que le racinisateur Snowball a opéré de faibles modifications morphologiques alors que le racinisateur pour l'arabe extrait la racine sémitique à 3 ou 4 lettres. Ainsi, la qualité des outils utilisés est aussi un paramètre à prendre en compte lors de l'analyse de l'apport des traitements associés.

## 5 Conclusion

Dans cet article, nous avons évalué l'impact des différents traitements linguistiques de base sur la tâche du résumé multi-documents multilingue. Les performances de notre système dépassent la Baseline. Les meilleurs résultats ont été enregistrés avec le filtrage des mots vides, des améliorations sont à prévoir afin de pénaliser les mots vides en utilisant mRMR.

Malgré les scores ROUGE pratiquement équivalents, le choix des phrases pour la construction des résumés varie considérablement d'une étape à une autre et ceci pour les trois langues. Cette variation s'explique par : la variation du score des formes que prend chaque terme à travers les différents niveaux d'analyse d'une part et par la qualité variable des outils utilisés pour chaque traitement d'une autre part. De plus, la variation des performances du système d'une langue à une autre est à étudier, les résultats pour l'arabe sont 10% moins bons que pour les autres langues, est ce que c'est uniquement du à la qualité des outils utilisés pour chaque langue ? Doit-on considérer la stabilité des performances à travers les langues comme critère d'évaluation ?

L'évaluation de l'efficacité du système et de la qualité des résumés produits à chaque étape nécessite une évaluation manuelle, les scores ROUGE étant pratiquement équivalents malgré la grande variation des choix de phrases fait resurgir de nouvelles questions. En effet, on constate de plus en plus que procéder à une évaluation automatique en utilisant ROUGE n'est pas suffisant. Elle doit être complétée par une évaluation manuelle pour juger des critères autre que la couverture lexicale : qualité linguistique, cohérence, etc.

## Références

- BARZILAY R., MCKEOWN K. R. (2005). Sentence Fusion for Multidocument News Summarization. *Comput. Linguist.*, 31(3), 297–328.
- BOUDIN F., TORRES-MORENO J.-M. (2009). Résumé automatique multi-document et indépendance de la langue: une première évaluation en français. Actes de *Traitement Automatique de La Langue Naturelle (TALN'09)*
- CONROY J., DAVIS S. T., KUBINA J., LIU Y.-K., O'LEARY D. P., SCHLESINGER J. D. (2013). Multilingual Summarization: Dimensionality Reduction and a Step Towards Optimal Term Coverage. Actes de *MultiLing 2013 Workshop on Multilingual Multi-document Summarization*, 55–63.
- CONROY J. M., SCHLESINGER J. D., KUBINA J., RANKEL P. A., O'LEARY, D. P. (2011). CLASSY 2011 at TAC: Guided and multi-lingual summaries and evaluation metrics. Actes de *Text Analysis Conference*.
- DAS P., SRIHARI R. (2011). Global and Local Models for Multi-Document Summarization. Actes de *Text Analysis Conference*.
- DING C., PENG H. (2003). Minimum Redundancy Feature Selection from Microarray Gene Expression Data. Actes de *IEEE Computer Society Conference on Bioinformatics (CSB'03)*, 523–529.
- GIANNAKOPOULOS G., EL-HAJ M., FAVR, B., LITVAK M., STEINBERGER J., VARMA V. (2011). TAC 2011 MultiLing Pilot Overview. Actes de *Text Analysis Conference (TAC2011)*.
- HMIDA F., FAVRE B. (2011). LIF at TAC Multiling: Towards a Truly Language Independent Summarizer. Actes de *Text Analysis Conference (TAC)*.
- LEDENEVA, Y., 2008. Effect of preprocessing on extractive summarization with maximal frequent sequences, *Actes de MICAI 2008: Advances in Artificial Intelligence*, 123–132.
- LIN, C. (2004). Rouge: A Package for Automatic Evaluation of Summaries. Actes de *Text Summarization Branches Out: ACL-04 Workshop*, 74–81.
- LITVAK M., LAST M., FRIEDMAN M. (2010). A New Approach to Improving Multilingual Summarization Using a Genetic Algorithm. Actes de *48th Annual Meeting of the Association for Computational Linguistics*, 927–936.

- MARCU D. C. (1998). The Rhetorical Parsing, Summarization, and Generation of Natural Language Texts. Actes de *35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*, 96-103.
- MIHALCEA R., TARAU P. (2005). A Language Independent Algorithm for Single and Multiple Document Summarization. Actes de *International Joint Conference on Natural Language Processing (IJCNLP)*, Vol. 5.
- PENG H., LONG F., DING C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27(8), 1226–1238.
- SAGGION H. (2011). Using SUMMA for Language Independent Summarization at TAC 2011. Actes de *Text Analysis Conference*.