A Fine-grained Multilingual Analysis Based on the Appraisal Theory: Application to Arabic and English Videos

Karima Abidi, Dominique Fohr, Denis Jouvet, David Langlois, Odile Mella, Kamel Smaïli

A Fine-grained Multilingual Analysis Based on the Appraisal Theory: Application to Arabic and English Videos

K. Abidi, D. Fohr D. Jouvet, D. Langlois, O. Mella, and K. Smaïli

Loria University of Lorraine, France
{abidi, fohr, jouvet, langlois, mella, smaili}@loria.fr

Abstract. The objective of this paper is to compare the opinions of two videos in two different languages. To do so, a fine-grained approach inspired from the appraisal theory is used to analyze the content of the videos that concern the same topic. In general, the methods devoted to sentiment analysis concern the study of the polarity of a text or an utterance. The appraisal approach goes further than the basic polarity sentiments and consider more detailed sentiments by covering additional attributes of opinions such as: Attitude, Graduation and Engagement.
In order to achieve such a comparison, in AMIS (Chist-Era project), we collected a corpus of 1503 Arabic and 1874 English videos. These videos need to be aligned in order to compare their contents, that is why we propose several methods to make them comparable. Then the best one is selected to align them and to constitute the data-set necessary for the fine-grained sentiment analysis.

Keywords: Video analysis · Sentiment analysis · Appraisal theory · Word embedding.

1 Introduction

The explosive growth of the communication tools such as the television and the Internet has facilitated the rapid broadcasting of the information. Consequently, several television programs and news are available in different languages. However, the access to the information expressed in a foreign language is inaccessible to many users. To tackle this problem, the AMIS (*Access to Multilingual Information and Opinions*) project proposes to develop a multilingual information comprehension help system without human intervention. AMIS is a Chist-Era project, the principal objective is to develop a system, helping people to understand the content of a source video by presenting its main ideas in a target understandable language. This system is based on several components such as: video summarization, audio summarization, text summarization, automatic speech recognition system, machine translation and sentiment analysis [23], [5] and [4]. Four architectures have been proposed, one of these scenarios is given in Fig.1, which corresponds to a pipeline assembly of some of the mentioned

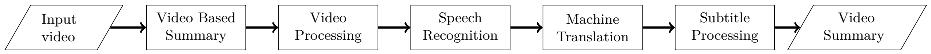

Fig. 1. Scenario 1 - the most basic approach to newscast summarization

components.
This architecture is the one that has been used in this article for our experiments.

Another aspect of AMIS is to compare two videos in two languages about the same topic and to produce a grain-fined sentiment analysis of their contents. In this article, we will focus only on this aspect of AMIS project. The rest of this paper is organised as follows. Section 2 presents the used video database. Then, we present an overview of the global model proposed to align and analyse the AMIS videos that deal with the same object in terms of opinions in Section 3. In Section 4, we describe the proposed method to identify the comparable AMIS videos. A fine-grained multilingual sentiment analysis approach is proposed in Section 5 and finally, we conclude.

2 Video Database of AMIS

In order to develop the AMIS system, a large corpus of newscasts and reports from different channels (see table 1) were crawled by using a list of controversial Hashtags (see table 2). More details on the crawling method is given in [14]. In Table 3, we give the number of the harvested videos for each monolingual corpora.

Table 1. The channels used for harvesting.

English channels	BBC news, France 24, RT, Euronews
Arabic channels	النهار, الشروق, العربية, القدس, الاولى Nessma, i24news, France 24, RT, Euronews, BBC news
French channels	France 24, RT, Euronews

Table 2. The used controversial Hashtags.

#Syria	#RealMadrid-FCBarcelona	#Animal-rights
#Trump	#Women's-rights	#Homosexual-marriage
#Drug-liberalization	#Death-sentence	#Occupied-territories

Table 3. The number of videos per language

Language	Number of videos
English	1874
Arabic	1503
French	2046

3 An overview of the global approach

Our objective is to make comparable the videos produced by the AMIS system. We have to mention that the comparability does not concern two well-written documents in the same language. In fact, we have two challenges to overcome in our case, the comparability is about the transcriptions of two speech recognition systems, one is in Arabic and the second is in English. That means that texts to make comparable include several errors. The second challenge concerns the multilingual aspect of the produced documents. In other words, we have to align two texts one is in Arabic and the second is in English. Several works, on multilingual comparability, have been proposed by the international community [10], [22], [9], [3], [13]. Overall, they concern documents harvested from social networks, Wikipedia, etc. But, in our knowledge there is very little work on the aspect of comparability on multilingual videos.

To achieve this goal, it is necessary to align the collected videos [12] and to take into account the qualitative aspect of the comparison material produced by the ASR systems. Once the comparable pairs are identified, the next step is to compare them in terms of opinions. In Fig.2, we present a global overview of the model we propose and that will be explained further.

Our approach is based on the use of two Automatic Speech Recognition systems (ASR), one in Arabic [18] and one in English [8]. It is also based on a module of videos alignment and an elaborated procedure of fine-grained sentiment analysis.

4 Identifying comparable videos

In this article, we will present two methods of comparability: one which is well widespread, it is based on a dictionary method, and a new one based on the word embedding (Word2vec) [19]. These methods are explained in detail in the following sections.

4.1 Dictionary-based method

The method consists in looking-up into a dictionary if the translation of the words of the source video V_s exist in the target one V_t and vice versa. The idea is to align all the pairs of videos that share as many words as possible between the source and the target videos. To do so, we need to measure the comparability

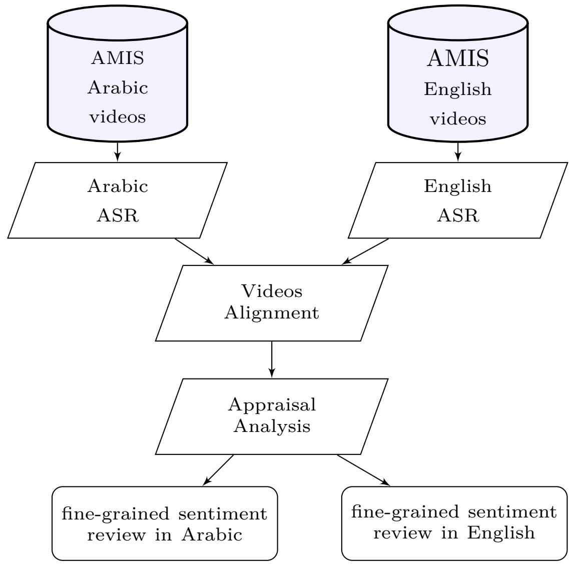

Fig. 2. An overview of the multilingual fine-grained sentiment analysis

between the videos pairs. The pair of videos that gets the best score is considered as the best comparable videos. For that, we used the well-known measure proposed by Li and Gaussier [15]. This comparability measure can be defined as the expectation of finding, for each English word w_e (respectively w_a) of the source video V_s (respectively of the target video V_t), its translation in the video V_t (respectively in the source video V_s).
The comparability measure is estimated as follows:

$$LG(V_s, V_t) = \frac{\sum_{w \in \{l_s \cap D_s\}} \sigma(w, l_t) + \sum_{w \in \{l_t \cap D_t\}} \sigma(w, l_s)}{\mid l_s \cap D_s| + \mid l_t \cap D_t|} \quad (1)$$

Where D_s is the source part (English) of the bilingual dictionary , D_t is the target part (Arabic) of the dictionary. l_s and l_t are respectively the list of words of the source and the target video.
σ is a function using two parameters: a word w and a list of words (l_s). This function indicates whether potential translations of the word w represented by the list $T(w)$ include at least one word in the list l_s.

$$\sigma(w, l_s) = \begin{cases} 1 \; if \; T(w) \cap l_s \neq \emptyset \\ 0 \; else \end{cases} \quad (2)$$

For this experimentation, we used the bilingual dictionary OMWN (Open Multilingual WordNet)[1] that contains 17,785 Arabic and English pairs.

The initial results of this approach led to bad performance. In fact, the drawback of this approach is its dependency of a bilingual dictionary. Whatever the size of this dictionary, the coverage issue arises especially for rich morphological language such as Arabic. In this language the word is composed, in the majority of cases, of the concatenation of a root and affixes. A root in Arabic is considered as a producer of words, that is why from a single root, several words can be produced. For example: the root كتب (*write*) with particular affixes produce different words with different meanings: يكتب (*he writes*), مكتبة (*library*), مكتب (*office*), etc. Consequently, in order to improve the coverage of the dictionary, we used the Buckwalter Arabic Morphological Analyzer to segment the words. Even if English does not have the same morphological constraints as Arabic, we also used a morphological analyzer (TreeTagger tool)[2] in order to reduce the missing inflected forms of words in the processed videos. In our experiments, the dictionary-based method includes in addition to the bilingual dictionary OMWN, all the inflectional form of its words.

As described above, the method necessitates a large bilingual dictionary, we replaced in another experiment the previous dictionary, by a translation table built on a parallel corpus of 9 million parallel sentences that led to a translation table of 297,176 pairs of Arabic and English entries [17].

4.2 Word embedding approach

The idea of this method is to investigate to what extent the semantic information encoded by words embedding approach can be used to retrieve the words semantically close to each other in two documents in which each of them is written in a different language. To do so, we used the CBOW method of Word2Vec model proposed by [19] to extract the bilingual vector representation of words. The CBOW method is trained over a large parallel corpus (9 million sentences in English and Arabic) with the objective to capture strong semantic relationships between the Arabic and English words. Each Arabic word is assigned a list of correlated English words which is calculated by a method proposed by the authors of [1].

To estimate the comparability between an Arabic and English videos, we used the same formula as in the previous section 4.1 except that σ is a function that returns 1 if a word in the target video exist in the correlated words list of a word of the source video.

[1] http://compling.hss.ntu.edu.sg/omw/

[2] `https://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/`

4.3 Experimentation

The methods presented previously are evaluated on a test corpus composed of 123 pairs of comparable videos extracted from the Euronews web site [8]. All the videos have been transcribed by our Arabic and English ASR systems depending on the language of the videos. The performance is given in terms of one of the classical measures in information retrieval topic: Recall (R@1, R@5 and R@10). In Table4, we report the results of the three methods: the one based on a dictionary with the use of the morphological analyzers ($DicMA$), the one based on a translation table ($DicTT$) and finally the one based on a bilingual vector representation of words ($CBOW$). This table shows that $DicTT$ achieves

Table 4. The performance of different comparability methods in terms of R@1, R@5, and R@10 on a test corpus.

Rappel	R@1	R@5	R@10
$DicMA$	43	65	76
$DicTT$	**70**	90	92
$CBOW$	39	62	75

the best results in comparison to the two others. The recall at rank 1 is 70% and grows up to 92% at rank 10. This result is encouraging, it allows, in almost cases to retrieve in the Top10 the right pair of comparable videos. The $CBOW$ method achieves similar result as the $DicMA$. This result is very interesting, since without external resources (a bilingual dictionary and a morphological analyzer), we can get almost the same performance. Consequently, this method could be used in under-resourced languages such as Arabic dialects.

By using the best method presented in this table, we retrieved all the pairs of comparable videos from the database of AMIS that led to 360 Arabic-English comparable videos. We recall that the total number of Arabic videos is 1,542, they concern several topics. Although videos were collected in different language using corresponding hashtags, that does not mean that each video in a given language has a matching comparable video in another language (in the collected video corpus). Furthermore, we selected only the pairs of videos for which the scores of comparability are high.

5 Multilingual Fine-granularity sentiment analysis

In general, the methods devoted to sentiment analysis concern the study of the polarity of a text or an utterance. The sentiments in this case are reduced to the three classical opinions: positive, negative or neutral. In some other studies, fine-grained categories are added to have a more detailed analysis by using emotions such as (anger, disgust, fear, joy, sadness, and surprise) [24] or by adopting a linguistic theory such as appraisal [25],[20],[11] and [2].

The appraisal approach has been developed by White and Martin [16] within the theory Systemic Functional Linguistics [6]. The idea is to go further than the basic polarity sentiments and consider more detailed sentiments by covering additional attributes of opinions such as: Attitude, Graduation and Engagement.

The theory is supported by a graph, which represents the different sentiment categories expressed by a speaker (Fig.3).

- **Attitude.** The category *Attitude* gives the type of appraisal being expressed as either *affect*, *appreciation*, or *judgment*.
  - *Affect.* This sub-category of *Attitude* describes the emotional reactions (happy, miserable, angry, etc.).
  - *Appreciation.* It concerns the opinion that a person has about the inner or outer qualities of an object (beautiful, innovative, amazing, etc.).
  - *Judgment.* This sub-category describes the behaviour of somebody in a social context (lucky, brave, famous, etc.).
- **Engagement.** Sentiment can be expressed directly or indirectly, it reflects the possibility of the production of an event (perhaps, seems, etc.)
- **Graduation.** This category refers to the strength or the force of emotion and attitude in each appraisal category. The graduation is globally expressed via modifiers, for example the combination of the 'modifier' "very" with an adjective intensifies the meaning of the utterance. There is another sub-category of Graduation, named *Focus*. It makes the meaning of something either more precise or less precise. For example: *a true challenge* or *it is a challenge*. In the first example, the challenge seems to be harder than in the second example.

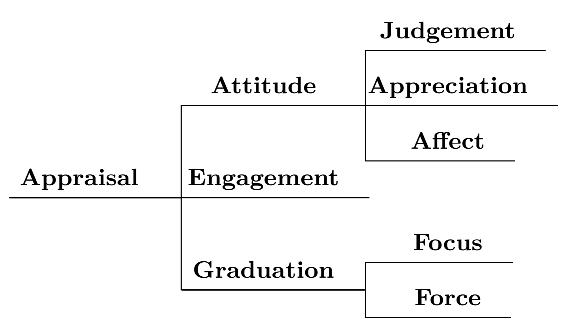

Fig. 3. Appraisal taxonomy

5.1 Building appraisal lexicon

In the following, for commodity reasons we will use only the categories: *Attitude* and *Graduation*. To build the appraisal lexicon, we started from an English opinion lexicon composed of 4,913 negative words and 2,718 positive words developed by Minqing Hu and Bing Liu [7]. Besides, we created a list of 363 words with their appraisal categories inspired from the examples of Martin and White's book [16] that we named *MW363*. This led to a list of words with their polarities and appraisal categories, some examples are given in Table 5. Then the idea is to

Table 5. Few examples of words with appraisal and polarity opinions

Word	Attitude sub-category	Polarity
Lucky	Judgment	Positive
Obscure	Judgment	Negative
Confident	Affect	Positive
Love	Affect	Positive
Helpful	Appreciation	Positive

use a lexicon with appraisal categories larger than the one we created (*MW363*). That is why, we decided to assign for each entry of the Bing Liu's lexicon the corresponding appraisal *Attitude* category by using a method combining Word2Vec and the $MW363$. The method consists in representing each word of respectively the Bing Liu's and $MW363$ lexicons by a word embedding approach by using the vectors trained on 100 billion words calculated from various news articles of Google[3]. To do so, for each word X from Bing Liu's lexicon, we find its top-n closest words to $MW363$. Each word of this list is labeled by a sub-category of *Attitude*. Then, we assign to X the sub-category which is predominant in this latter list.
Since the words of Bing Liu's lexicon have already polarity signs, when we assign them an appraisal sub-category, we get new sub-category with a polarity. That means, for example, a word may have an *Affect* sub-category but this one will be signed by the initial polarity. Each word of the Bing Liu's lexicon will be assigned an appraisal positive or negative score (S_{app}) calculated as in the formula 3. A positive or negative score respectively indicates how positive or negative is the word in terms of the Attitude sub-category. The achieved lexicon that corresponds to the initial lexicon of Bing Liu is henceforth increased by the *Attitude* appraisal category. It will be referred in the following as BingApp.

$$S_{App}(X) = \frac{1}{d_n} \sum_{i=1}^{d_n} cosine(X, W_i) * P_{W_i} \tag{3}$$

Where:

[3] `https://code.google.com/archive/p/word2vec/`

- d_n : The number of words in the predominant sub-category in the list of the n closest words with X.
- W_i : A word belonging to the list of the predominant attitude sub-category.
- X : A word of Bing Liu's sentiment lexicon.
- $P_{W_i} = \begin{cases} +1 \text{ if } W_i \text{ is positive} \\ -1 \text{ otherwise.} \end{cases}$

We recall that our objective is to compare two videos one in English and the other in Arabic in terms of fine-grained opinions. In order to work with the same material in Arabic and in English, we translated BingApp into Arabic and we kept for each Arabic word the same sub-category and the same score as the English word. In Table 6 we give few examples of the achieved lexicon.

Table 6. Few examples of BingApp.

English word	**Arabic translation**	**Appraisal categories /Sub-category**	S_{App}	**Polarity**
Criminal	مجرم	Attitude/Judgment	-0.45	N
Attentive	منتبه	Attitude/Judgment	0.41	P
Worried	قلق	Attitude/Affect	-0.45	N
Satisfied	راض	Attitude/Affect	0.24	P
Harmonious	متناغم	Attitude/Appreciation	0.63	P

5.2 Fine-granularity sentiment predicting model

To be able to make an efficient fine-grained sentiment analysis, we need to enrich BingApp by adding other categories. To do so, we have to take into account, at least, two linguistic phenomena. To illustrate our purpose, let study the following example: *This cake is not very good.* We can remark that this sentence contains a negation form that precedes the phrase (*very good*). Consequently, the underlying opinion of this sentence can be completely inverted.
In this example, the adverb (*very*) is used to emphasis the adjective *good.* In other words, it modifies its intensity by adding force to this adjective. This phenomenon must be considered, especially knowing that the *Force* is an existing sub-category of the *Graduation* category.

- *Dealing with the Negation.* We added a new category to the appraisal taxonomy that we called *Inversion*, with its sub-category *Negation.* Then, we added to BingApp, the negation words (*Not, No, Neither, Nor, etc.*) and assigned them to the *Inversion* category. During the analysis step, if the *Inversion* category is identified in an utterance, then the polarity of the word following the negation item is inverted.
- *Dealing with the Force.* To consider the *Force* in the analysis, we added to the dictionary BingApp several modifier words that we assigned to the

sub-category *Force* of the category *Graduation*. We shared these modifiers through 4 classes. Each class indicates the intensity of the modifier and it is assigned a score proportional to its capacity to intensify a word. These weights have been set by hand. In Table 7, we give some examples of the new *Force* classes and their corresponding words that have been inserted into BingApp.

Table 7. The four classes of the intensity modifiers.

Force classes	Modifiers
Extreme	hardly, scarcely, barely, very, greatly, etc.
High	large, less, distant, more, etc.
Moderate	somewhat, relatively, rather, reasonably, many, etc.
Low	slightly, least, small, etc.

In Fig.4, we illustrate the new taxonomy of the appraisal theory.

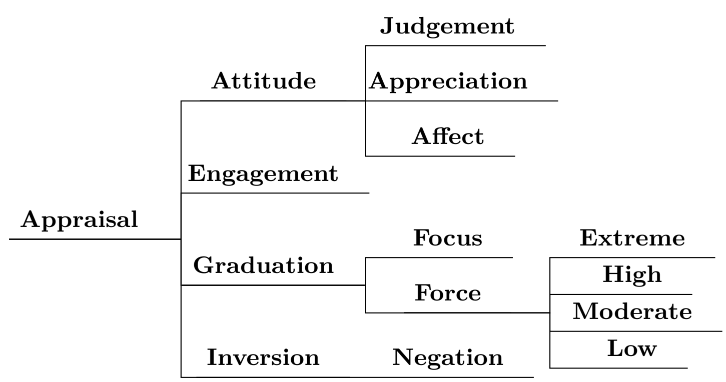

Fig. 4. Appraisal taxonomy

In order to evaluate the quality of our lexicon, we decided to use it in the assessment of the publicly available collection of movie reviews constructed by [21]. This standard test consists of 1,000 positive and 1,000 negative reviews. In order to study the impact of the use of the Inversion category that we added to the appraisal theory, we selected only the reviews that are concerned by this category. This led to a test corpus of 992 reviews including 538 positive reviews. In Table 8 we reported the recall and the precision values obtained by using the standard Bing Liu's lexicon and by the lexicon we created BingApp.

Table 8. Comparison of Bing Liu's lexicon and BingApp on a test Review corpus

Method	Recall	Precision
Bing Liu's	69.0	68.8
BingApp	**70.9**	**71.0**

In this experiment BingApp yields to better performance even if the difference is not very important. This test is not the main result of this work. It has been done only in order to know whether we use Bing Liu's or BingApp lexicon for evaluating the opinions underlying the videos. In conclusion, we consider that the appraisal approach led to better results, thanks to the dictionary we created, in comparison to the classical method based on the polarity supported by Bing Liu's lexicon.

5.3 Evaluation on AMIS videos

In the following, we propose to assess finely the opinions within the videos by using BingApp, the appraisal lexicon we created. A quantitative and qualitative evaluation is proposed. Each video is evaluated by a score we propose in formula 4.

$$S = \sum_{i=1}^{N} \alpha(w_{i-k}^{i-1}) * S_{App}(w_i) \tag{4}$$

Where N is the size of the video in terms of number of words. α is a weight depending on the *Inverted* or the *Force* sub-category of the k words preceding the word w_i (k is set to 2). It is the size of the cache in which the *Force* or the *Negation* are looked for.

The second assessment focuses on a qualitative evaluation in which we summarize the expressed opinion in the video. The idea is to facilitate the interpretation of the underlying opinion within a video and not just give an overall assessment score. A template of the opinion review is proposed in Fig.5.

The sentiment of the video is positive with a score $[X_p]$ and negative with a score $[X_n]$. X_{aff}% of the video concerns emotional reactions. X_{Jug}% of the video concerns the human behaviour according to social norms and $1-(X_{aff}+X_{Jug})$% of the video is about the appreciation of no human being entities. The force of the subject: ([target word]) is [augmented/reduced] thanks to the word [Modifier]

Fig. 5. The template used to generate the qualitative evaluation

This template corresponding to the review presented to the user indicates how much the video is negative or positive? What is the percentage of each sub-category of the category *Attitude*? Which word has been augmented or reduced? And which word participated to the augmentation or the reduction of the *Force*. An example is given in Fig.6.

Example: *well, to coin a phrase, the reports of "babe: pig in the city" 's death at the hands of a dark, scary, Felliniesque interpretation have been greatly exaggerated.*
Evaluation
The sentiment of the video is positive with a score **0.45** and negative with a score**-2.28**. **22.22**% of the video concerns emotional reactions. **44.44**% of the video concerns the human behaviour according to social norms and **33.34**% of the video is about the appreciation of no human being entities. The force of the subject: (**interpretation**) is **augmented** thanks to the word **greatly**

Fig. 6. An example of qualitative evaluation.

6 Conclusion

In this article the objective was twofold. The first one consisted in aligning the videos of AMIS project by making comparable the Arabic and the English videos describing the same subject . We tested three methods and compared them. The best one has been used to align the whole database of AMIS. Then we used a new method based on the appraisal approach allowing to have a fine-grained opinion analysis. For that, we created a new lexicon including more than 7,000 entries, each of them is assigned to the appraisal category. This dictionary served to evaluate quantitatively and qualitatively the content of videos. A review template has been proposed to summarize the opinions inside the video.

7 Acknowledgements

We would like to acknowledge the support of Chist-Era for funding this work through the AMIS (Access Multilingual Information opinionS) project.

References

1. Abidi, K., Menacer, M.A., Smaïli, K.: CALYOU: a Comparable spoken ALgerian corpus extracted from YOUTube. In: 18^{th} Annual Conference of the International Speech Communication Association, Stockholm Sweden, Interspeech (2017)
2. Alamsyah, A., Rahmah, W., Irawan, H.: Sentiment analysis based on appraisal theory for marketing intelligence in indonesia's mobile phone market (2015)

3. Barrón-Cedeño, A., España-Bonet, C., Boldoba, J., Màrquez, L.: A factory of comparable corpora from wikipedia. In: Proceedings of the Eighth Workshop on Building and Using Comparable Corpora, Beijing, China, July 30, 2015. pp. 3–13 (2015)
4. Garcia-Zapirain, B., Castillo, C., Badiola, A., Zahia, S., Mendez, A., Langlois, D., Jouvet, D., Torres-Moreno, J.M., Leszczuk, M., Smaïli, K.: A Proposed Methodology for Subjective Evaluation of Video and Text Summarization. In: MISSI 2018 - 11th edition of the International Conference on Multimedia and Network Information Systems. Advances in Intelligent Systems and Computing, vol. 833, pp. 396–404. Springer, Wroclaw, Poland (2018)
5. Grega, M.L., Smaïli, K., Leszczuk, M., González-Gallardo, C.E., Torres-Moreno, J.M., Linhares Pontes, E., Fohr, D., Mella, O., Menacer, M.A., Jouvet, D.: An Integrated AMIS Prototype for Automated Summarization and Translation of Newscasts and Reports. In: MISSI 2018 - 11th International Conference on Multimedia and Network Information Systems. vol. 833, pp. 415–423. Springer, Wroclaw, Poland (2018)
6. Halliday, M.A.K.: An Introduction to Functional Grammar. Hodder Arnold, third edition (1994)
7. Hu, M., Liu, B.: Mining and summarizing customer reviews. In: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '04 (2004)
8. Jouvet, D., Langlois, D., Menacer, M.A., Fohr, D., Mella, O., Smaïli, K.: About vocabulary adaptation for automatic speech recognition of video data. In: ICNLSSP'2017 - International Conference on Natural Language, Signal and Speech Processing. pp. 1–5. Casablanca, Morocco (2017)
9. Karima, A., Smaïli, K.: Measuring the comparability of multilingual corpora extracted from Twitter and others. In: series, S.L. (ed.) The Tenth International Conference on Natural Language Processing (HrTAL2016). Dubrovnik, Croatia (Sep 2016)
10. Kim, J., Li, J.J., Lee, J.H.: Evaluating multilanguage-comparability of subjectivity analysis systems. In: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. pp. 595–603. ACL '10, Association for Computational Linguistics, Stroudsburg, PA, USA (2010)
11. Korenek, P., Simko, M.: Sentiment analysis on microblog utilizing appraisal theory. World Wide Web **17**(4), 847–867 (2014)
12. Koźbiał, A., Leszczuk, M.: Collection, analysis and summarization of video content. In: Choroś, K., Kopel, M., Kukla, E., Siemiński, A. (eds.) Multimedia and Network Information Systems. pp. 405–414. Springer International Publishing, Cham (2019)
13. Langlois, D., Saad, M., Smaïli, K.: Alignment of comparable documents: comparison of similarity measures on French-English-Arabic data. Natural Language Engineering (Jun 2018). https://doi.org/10.1017/S1351324918000232
14. Leszczuk, M., Grega, M., Koźbiał, A., Gliwski, J., Wasieczko, K., Smaïli, K.: Video summarization framework for newscasts and reports – work in progress. In: Dziech, A., Czyżewski, A. (eds.) Multimedia Communications, Services and Security (2017)
15. Li, B., Gaussier, É.: Improving corpus comparability for bilingual lexicon extraction from comparable corpora. In: COLING 2010, 23rd International Conference on Computational Linguistics, Proceedings of the Conference, 23-27 August 2010, Beijing, China. pp. 644–652 (2010)
16. Martin, J., White, P.: The Language of Evaluation Appraisal in English. Palgrave Macmillan, London (2005)

17. Menacer, M.A., Langlois, D., Mella, O., Fohr, D., Jouvet, D., Smaïli, K.: Is statistical machine translation approach dead? In: ICNLSSP 2017 - International Conference on Natural Language, Signal and Speech Processing. pp. 1–5. ISGA, Casablanca, Morocco (2017)
18. Menacer, M.A., Mella, O., Fohr, D., Jouvet, D., Langlois, D., Smaili, K.: An enhanced automatic speech recognition system for arabic. In: Proceedings of the Third Arabic Natural Language Processing Workshop. pp. 157–165 (2017)
19. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. In: ICLR (Workshop) (2013)
20. Momtazi, S.: Fine-grained german sentiment analysis on social media. In: LREC (2012)
21. Pang, B., Lee, L.: A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In: Proceedings of the ACL (2004)
22. Saad, M., Langlois, D., Smaïli, K.: Extracting comparable articles from wikipedia and measuring their comparabilities. Procedia - Social and Behavioral Sciences **95**, 40 – 47 (2013)
23. Smaïli, K., Fohr, D., González-Gallardo, C., Grega, M., Janowski, L., Jouvet, D., Komorowski, A., Kozbial, A., Langlois, D., Leszczuk, M., Mella, O., Menacer, M.A., Mendez, A., Linhares Pontes, E., Sanjuan, E., Swist, D., Torres-Moreno, J.M., Garcia-Zapirain, B.: A First Summarization System of a Video in a Target Language. In: Springer Proceedings MISSI 2018 (2018)
24. Strapparava, C., Mihalcea, R.: Semeval-2007 task 14: Affective text. In: Proceedings of the 4th International Workshop on Semantic Evaluations. pp. 70–74. SemEval '07, Association for Computational Linguistics, Stroudsburg, PA, USA (2007)
25. Whitelaw, C., Garg, N., Argamon, S.: Using appraisal groups for sentiment analysis. In: Proceedings of the 14th ACM International Conference on Information and Knowledge Management. CIKM '05 (2005)