EEA Financial Mechanism 2009-2014

XENO@GR

Deliverable D4.3

Report on the Knowledge Network

Project Number: EOX GR07/3712

Project Title: Examining the phenomenon of xenophobia in Greece during the economic crisis: A computational perspective

Subproject 2: Computational methods and techniques for analyzing xenophobia in Greece (ΣΑΕ”013/8”)

ICELAND LIECHTENSTEIN NORWAY eea grants

ATHENA Research & Innovation Information Technologies

ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ, ΕΡΕΥΝΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ
ΓΕΝΙΚΗ ΓΡΑΜΜΑΤΕΙΑ ΕΡΕΥΝΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ

ICELAND LIECHTENSTEIN NORWAY eea grants

ATHENA Research & Innovation Information Technologies

Deliverable 4.3
Report on the Knowledge Network

Date:	December 2016
Authors:	Maria Pontiki, Dr. Haris Papageorgiou, Konstantina Papanikolaou, Dimitris Pappas, Dimitris Gkoumas, Penny Labropoulou, Maria Gavriilidou, Stelios Piperidis
Dissemination level:	Public
Version:	1.0
Keywords:	Social Media, Sentiment Analysis, Verbal Aggressiveness, Xenophobia
Description:	This deliverable presents the methodology followed for building a Knowledge Network of verbal aggressive Twitter messages towards predefined target groups of interest. The Knowledge Network enables to monitor and measure Verbal Aggressiveness as an indicator of xenophobic attitudes providing insights for the research questions the current project aims to address.

Table of Contents

Table of Figures

Table of Tables

1. Introduction

The current project aims to examine the phenomenon of xenophobia in Greece through a large-scale multi-source study based on the use of advanced computational social science and text mining approaches. The phenomenon of xenophobia is often examined from the social sciences point of view with more traditional in nature tools (i.e. interviews, surveys, analysis of secondary literature, questionnaires). Taking advantage of computational techniques and text mining methods facilitates the collection and the analysis of massive amounts of useful and unexplored data. For example, with the advent of Social Media (e.g. Twitter) and online fora people publicly voice their sentiments and beliefs without being asked to do so; such data are freely available in massive amounts providing new paths for political and social science research.

The ongoing economic and refugee/immigrant crisis in Europe gave burst to anti-immigrant sentiments, attitudes and practices across Europe ranging from individual (re)actions to official state policies (e.g. closing borders). Focusing on the Greek case, the main research puzzle in this project is whether (or not) the phenomenon of xenophobia is an outcome of the recent financial crisis or it comprises a long-lasting social perception deeply rooted in the Greek society. A core activity towards addressing these research questions is to explore current aspects of the phenomenon of xenophobia with a focus on the discursive practices employed in everyday discourse and in particular to harvest, exploit and incorporate knowledge from Social Media channels.

Xenophobia as a “psychological state of hostility or fear towards outsiders” (Reynolds and Vine, 1987) is associated with feelings of dominance (implying superiority) or vulnerability (implying the perception of threat), respectively (van der Veer, 2013). Focusing on Verbal Aggressiveness (VA) as an important component of the manifestations of xenophobia expressing feelings of vulnerability or/and feelings of superiority towards those perceived as “foreigners”, we have designed and built a knowledge network of on line expression of VA towards specific target groups of interest (e.g. Jews, Muslims, Albanians, etc.).

The knowledge network-database helps to study the formulation of VA in relation to specific target groups, to measure and monitor different aspects of VA in time and provides insights for the research questions the particular project aims to address.

To this end, the main scope of this deliverable is to present the methodology followed for building and populating the knowledge database with the output of the automatic VA analysis of Twitter user generated content.

2. Methodology

The overall workflow for building the knowledge network is a 6-step process presented below in Figure 1. The first step was to gather data related to specific target groups (TGs) of interest (e.g. Jews, Muslims, Albanians, etc.). The TGs -10 in total- were defined based on a number of criteria (e.g. population of the specific ethnic groups in Greece, dominant prejudices in Greece about the specific groups). In a second phase, samples of the collected data were explored by experts in order to identify different aspects of VA related to the predefined entities of interest. Based on data observations and literature review we, then, designed a linguistically-driven VA framework according to which the VA messages (VAMs) are classified into distinct categories based on specific semantic criteria (described below in section 2.3). The next step was the design and the development of the resources (e.g. lexical resources, linguistic patterns) and the models/algorithms needed for the computational treatment of the VA framework (VA analyzer). Subsequently, the data collections were automatically processed using the VA analyzer and the Knowledge Database was populated with the output of the VA framework. Finally, the output was visualized in various ways in order to obtain a better understanding of the data and the results of the VA analysis.

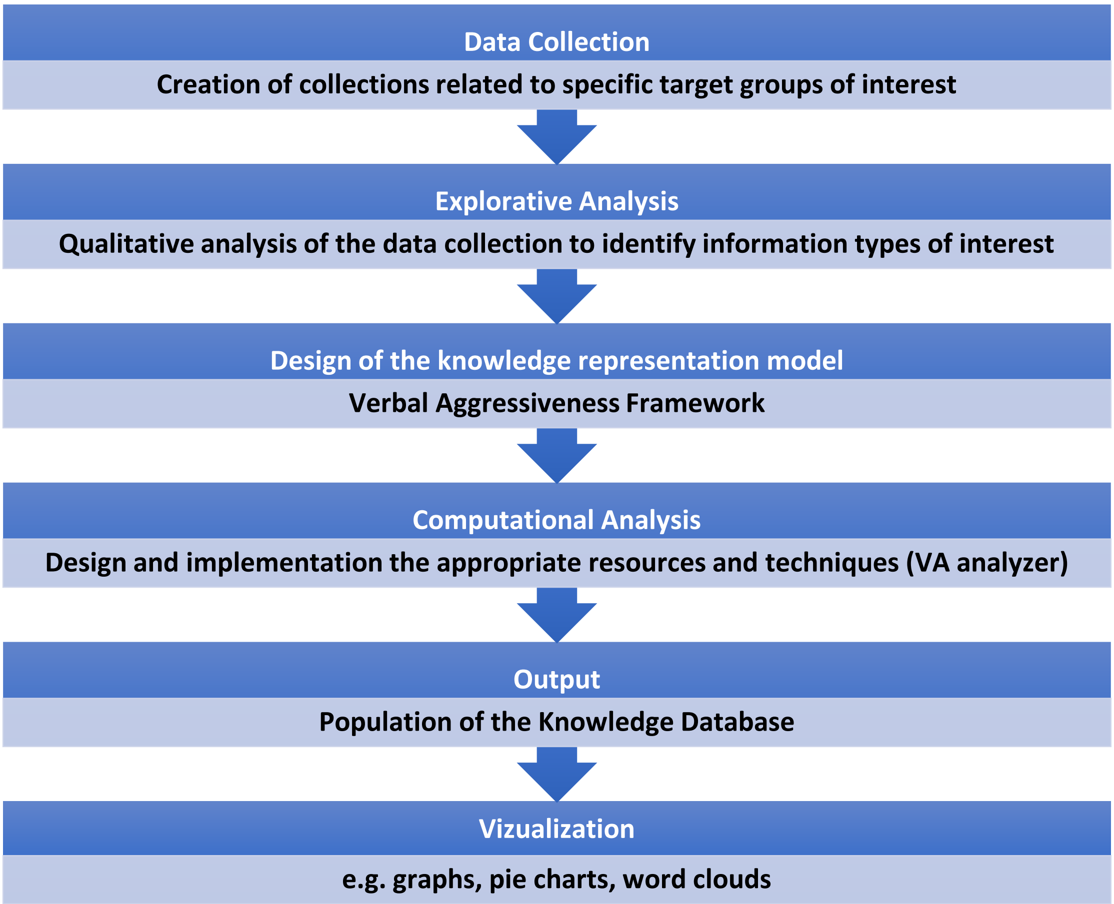

Figure 1: Workflow for building the knowledge network

2.1 Data Collection

For each TG of interest we retrieved from the Twitter data source relevant tweets using related queries/keywords (e.g. “ισλάμ” (=“islam”), “Πακιστανός” (=“Pakistani”), “Ρουμάνος” (=“Romanian”), etc.) covering the time period 2013-2016. Given that the search function in the database configuration is stemmed, the queries returned also tweets containing morphological variations of the selected keywords (e.g. “*ισλαμοποίηση*” for “*ισλάμ*”); the search resulted in 10 collections (1 per TG) containing in total **4.490.572** Tweets (see Table 1).

Target Group	Keyword(s)/Queries	Number of Tweets
TG1: Pakistani	Πακιστανός	**66.307**
TG2: Albanians	Αλβανός	**199.095**
TG3: Romanians	Ρουμάνος	**74.270**
TG4: Syrians	Σύρος	**299.350**
TG5: Muslims	Μουσουλμάνος, ισλάμ	**546.880**
TG6: Jews	Εβραίος	**101.262**

TG7: Germans	Γερμανός	**1.097.597**
TG8: Roma	Τσιγγάνος	**182.974**
TG9: Immigrants	Μετανάστης, αλλοδαπός	**672.009**
TG0: Refugees	Πρόσφυγας	**1.250.828**

Table 1: Data collection per TG

The per-year amount of Tweets that were retrieved for each TG is illustrated below in Figure 2. The most discussed TG is the one of the "refugees". In particular, the amount of Tweets mentioning "refugees" has been rapidly increased during the refugee crisis in Europe (2015 and 2016). The second most discussed TG is the one of "Germans", with "Immigrants" and "Syrians" following third and fourth, respectively.

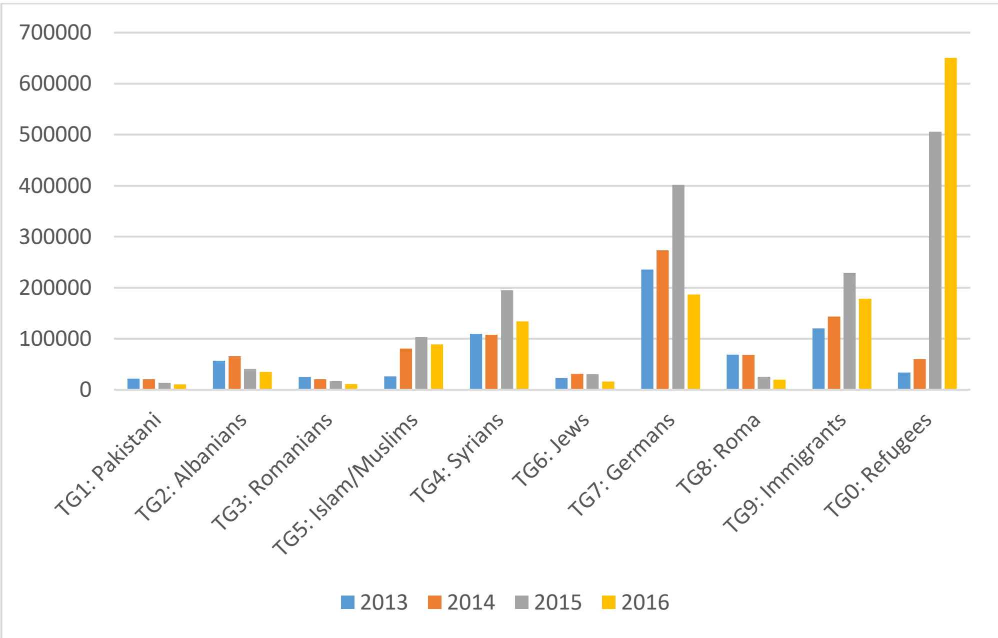

Figure 2: Data collections per year

2.2 Data Exploration

Data exploration is an integral part of the methodology, since it helps to understand and obtain a broader view of the whole dataset and is crucial for filtering the data and clustering them into targeted collections that can be used for development and training purposes.

To this end samples of the collected data were explored by experts (computational linguists and political scientists) using the Palomar Data Analysis and Modeling

Platform (Papanikolaou et al., 2016). In particular, the Tweets were examined from two different yet interconnected perspectives:

- Focusing on the types of the verbal attacks (i.e. different aspects of VA) against the TGs as well as on the types of linguistic weapons used for the attacks (i.e. linguistic instantiations of VAMs).

- Focusing on the emerging stereotypes and themes discussed per TG (e.g. criminality stereotype, personal hygiene stereotype, cultural inferiority etc.).

This was an iterative procedure, as simultaneously the VA Framework was modified and improved, until it was finalized according to the exploration. The outcome of this phase was VA oriented data collections that were used for the development of the VA analyzer.

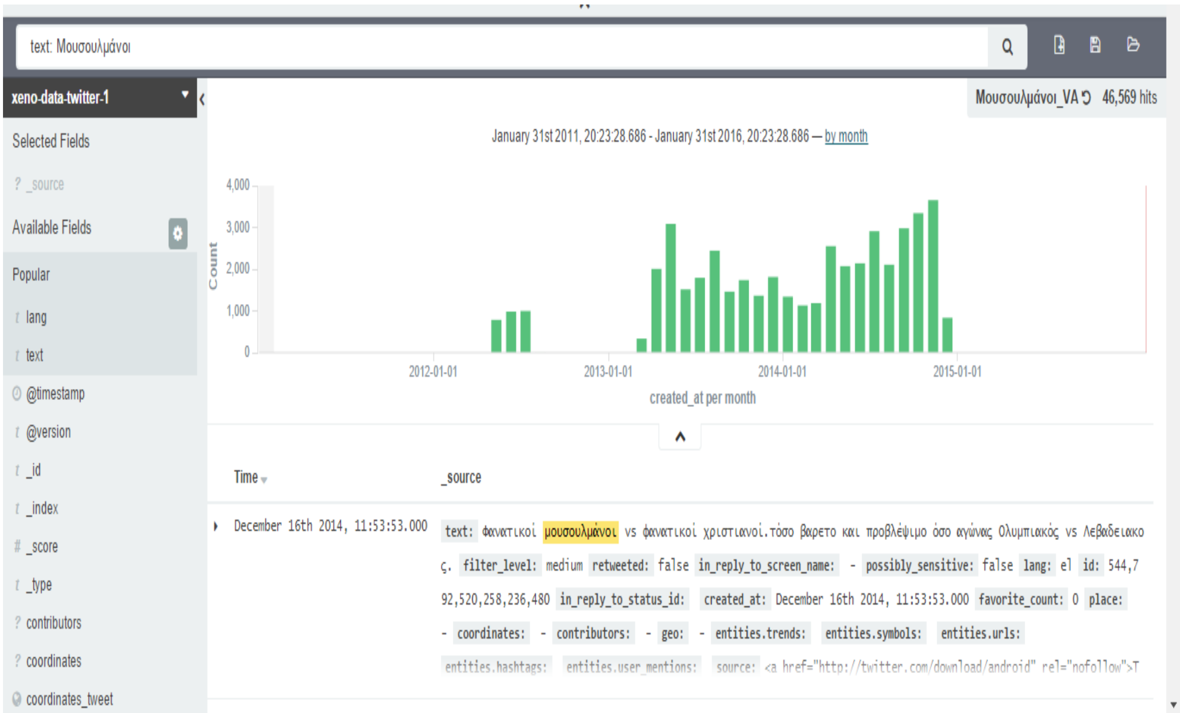

Figure 3: Exploring the twitter collection, retrieving documents using the query "Muslim"

2.3 Verbal Aggressiveness as an indicator of xenophobic attitudes

2.3.1 Basic concepts, definitions and typologies

Verbal Aggression comes about as a part of hostility which is an intrinsic aspect of personality (Infante, 1987), and involves using messages to attack other people or those aspects of their lives that are extensions of their identity" (Hamilton and Hamble, 2011). The concept and the content of Verbal Aggressiveness (VA) has been studied within the scope of psychology and communication studies (Hamilton and Hamble, 2011; Infante and Wigley, 1986; Kinney, 1994) in different contexts (e.g. marriage, workplace, parental relations).

Online VA can be termed generally as "flaming" including everything from impoliteness and swearing to excessive use of exclamations and superlatives (Kiesler et al., 1984) and it is often examined with reference to xenophobia and racism (Laineste, 2012). Flaming has come to be seen as a common term to designate any negative and antisocial verbal behavior on computer networks – e.g. as a "form of personal verbal violence arising largely from the peculiar conditions of online writing" (Millard, 1997; Tereszkiewicz 2012). VA case studies have provided some useful insights on the use of the internet and the on line expression of aggressiveness often with reference to expressions of xenophobia and racism. The internet and on line communication have increased the expression of verbal aggressiveness. Anonymity and "volatile identities" are among the factors that account for this trend, though not exclusively.

Depending on the approach, the social context (e.g. marriage, workplace, etc.) and the communication type (face to face vs online communication) several typologies of VA messages (VAMs) have been proposed. Infante (1987) and Infante et al (1990) suggest a ten-way classification schema:

- Character attack
- Competence attack
- Background attack
- Physical appearance attack

- Malediction
- Teasing
- Ridicule
- Threats
- Swearing
- Nonverbal emblems

Kinney (1994) suggests a typology of verbal aggression based on the domains of attacks. In particular, his typology involves the following:

- **Group membership attacks** (messages that associated or placed one into a negatively evaluated group).
- **Personal failings attacks** (messages that pointed out personal deficits).
- **Relational failings attacks** (messages that described one's social or interpersonal relationship deficits).

According to Kinney (1994) there is correspondence with the classification schema of Infante et al. (1990) involving background attacks, character attacks, competence attacks, and physical appearance attacks. The fact that maledictions, teases, ridicules, threats and swears did not surface in the current results suggests that they may represent methods of attack rather that targets of attack.

The VA typology for online contents (Laineste, 2012; Verkhovsky, 2006) considers three types of aggression based on its intensity as follows:

- **Strong** aggression: when a text expresses straightforward violence, displays nationalist or racist slogans, calls for physical actions against "others", and praises historical violence.
- **Medium** aggression: uses or introduces new negative stereotypes about the "other", swearing, accusing of stupidity, naming and slurs.
- **Mild** aggression: when jokes and other forms of humour are used, the target is presented in a negative context, or as possessing negative influence, racist viewpoints are referred to or a previous flame is cited without any counter-arguments.

Profanity and physical threats are perceived as more aggressive than criticism (Greenberg, 1976). However, depending on the social context the intensity of aggression may vary. For example, swearing may indicate high (Infante et al., 1992) or medium aggression (Laineste, 2012). In addition, the data source also matters when classifying aggression focusing on intensity, since different online contexts (e.g. social media, hate-promoting websites, etc.) do differ in the level of aggression (Laineste, 2012).

2.3.2 VA Framework

Based on literature review and explorative analysis findings we propose a linguistically-driven VA framework where VA messages (VAMs) are classified into distinct categories based on specific linguistic criteria. The concept of VA presupposes the speech act theory performative approach to language, which addresses speaking as intentionally doing things with words (Austin, 1962). Moving a step forward in the perception of verbal aggression, the intentional use of language can be associated with the social construction of aggression; thus, in terms of social psychology, language can be viewed as a "weapon" (Graumann, 1998).

We employ a data-driven approach focusing on explicitly stated aggressive messages/expressions towards the TGs of interest. Given a collection of Tweets, the goal is to identify different types of verbal attacks against the TGs following the typology presented below in Figure 4.

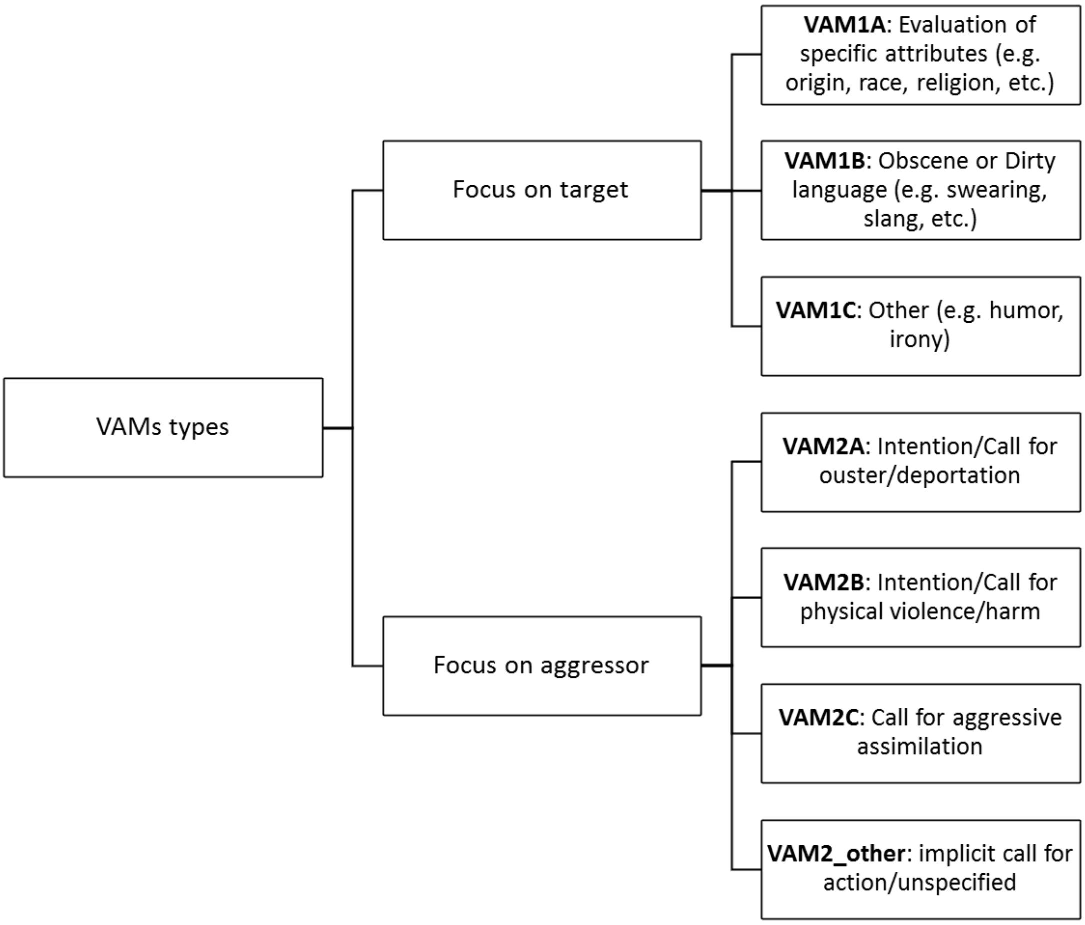

Figure 4: Typology of VAMs

As illustrated above in Figure 4, VA messages (VAMs) are classified into distinct categories based on:

- Their focus (i.e. distinguishing between VA utterances focusing on the target of the attack and VA utterances focusing on the attacker).
- The type of linguistic weapon used for the attack (e.g. formal evaluations, obscene/dirty language, humor).
- The content of the attack (e.g. threats/calls for physical violence or for deportation).

In particular, we consider two main types of VAMs (VAM1 and VAM2) that are further categorized in specific subtypes:

- **VAM1:** Messages of this type focus on (the attributes of) the target (e.g. physical appearance, religion, etc.) and are further classified into subcategories based on the type of the linguistic devices (weapons) used by the aggressor to attack the target:

- **VAM1A**: Formal evaluation of specific attributes (e.g. origin, race, religion, etc.) e.g.
  "και κάτι που ξέχασα να προσθέσω είναι ότι η θρησκεία (Μουσουλμάνοι) δεν χαρακτηρίζεται από καινοτομίες..."
  [*"Islam is not characterized by innovation (= meaning forward thinking)…"*]

- **VAM1B**: Taboo or dirty language (e.g. swearing, slang, etc.) e.g.
  " Γαμω τους αλβανούς ρε φίλε...."
  [*"Fucking Albanians…"*]

Note that messages of this type may also express evaluation about specific attributes (e.g. dimwit). Obscene messages are considered a separate category because they can provide different types of insights. For example, as mentioned above in section 2.3.1, depending on the online context, swearing may indicate different levels of aggression. In addition, swearing can act as an in-group solidarity marker, as when a group shares identical swearing norms (Mercury 1995; Allan and Burridge, 2006; Crystal 1995).

- **VAM1C**: Other (e.g. humor, irony) e.g.
  "Ευτυχώς που η φύση κρατάει ισορροπία και σκοτώνονται οι Εβραίοι με τους φανατικούς μουσουλμάνους !!!"
  [*"Jews and Muslims are killing each other…fortunately nature keeps a balance!!!"*]

- **VAM2:** Messages of this type focus on the aggressor's intentions providing information about specific types of attack and are further classified into subcategories based on content the of the attack:

- **VAM2A**: Intention or call for ouster/deportation (oriented to legal means) e.g.
  "Άμεση απέλαση... Αφού δεν σέβονταν την χώρα... RT @skaigr Εξέγερση μεταναστών στο κέντρο της Αμυγδαλέζας..."
  [*"Immediately deport the immigrants…they do not respect our country"*]

- **VAM2B**: Intention or call for physical violence/harm (oriented to physical extinction) e.g.
  "Οι φυλακισμενοι, οι φιλοι μου να εκτελεσουν την παραγγελια..... ΦΡΙΚΤΟΣ θανατος στο Πακιστανικο κτηνος"
  [*"Murder that Pakistani beast"*]

- **VAM2C**: Call for aggressive assimilation e.g.
  "Να εκχριστιανιστούν οι Μουσουλμάνοι μετανάστες αν θέλουν άδεια εργασίας στην Ελλάδα. Μαθήματα γλώσσας κι ελληνικής ιστορίας."
  [*"Muslims should be baptized if they want to find a job in Greece"*]

- **VAM2_other**: Implicit or unspecified call for action e.g.
  "Θα συνεχίσουμε να κάνουμε τους χαζούς μπροστά στον ισλαμικό κίνδυνο;"
  [*"We will keep pretending that there is no Islamic danger?"*]

2.4 Verbal Aggressiveness Analysis

For the computational treatment of the proposed framework we have designed a linguistically-driven VA analyser that given an input text (i.e. a Tweet) detects VAMs towards the TGs of interest and classifies them according to the typology presented in section 2.3. In particular, the output of the VA analyser contains the following 4 types of information:

- **TG_id**: the unique ID number that has been assigned for each TG of interest (predefined values: TG0-TG9, see Table1, section 2.1)
- **TG_evidence**: The lexicalization of the TG as referred to in the Tweet.

- **VAM_type**: the type of the VAM as it is coded in the typology (predefined values: VAM1A, VAM1B, VAM1C, VAM2A, VAM2B, VAM2C, VAM2other)
- **VA_evidence**: The lexicalization of the verbal attack as it appears in the Tweet.

For example, for the Tweet:

> *"Να <u>εκχριστιανιστούν</u> οι Μουσουλμάνοι μετανάστες αν θέλουν άδεια εργασίας στην Ελλάδα. Μαθήματα γλώσσας κι ελληνικής ιστορίας."*
> [*"Muslims should be <u>baptized</u> if they want to find a job in Greece"*]

the VA analyser returns the following tuple:

[TG_id: "TG5", TG_evidence: "*Μουσουλμάνοι*", VAM_type: "VAM2C", VA_evidence: "*εκχριστιανιστούν*"]

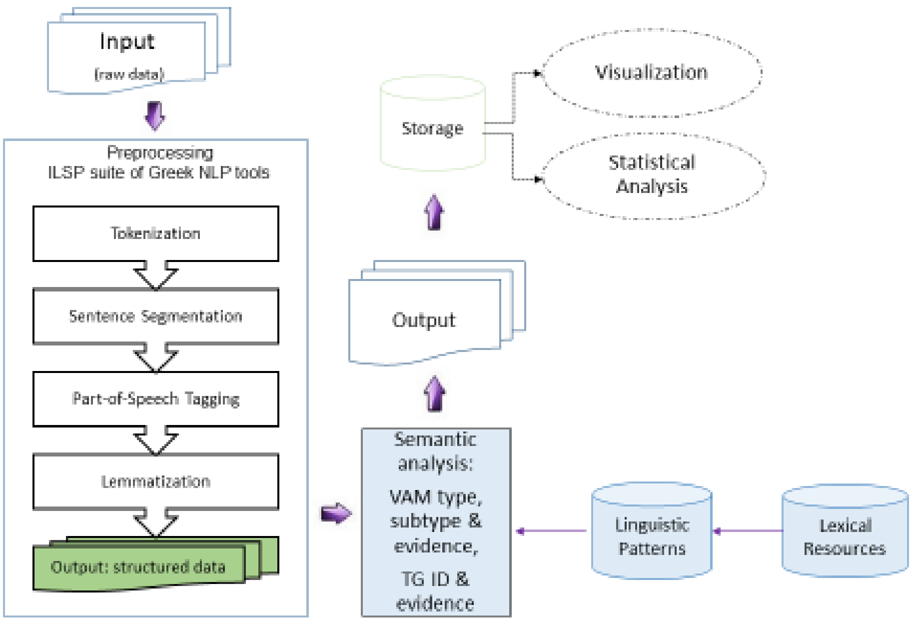

Figure 5: Architecture for VA analysis

The overall architecture for the VA analysis is illustrated in Figure 5. The input for the VA analyser is raw data (Twitter collections). In a first phase the data was processed through a Natural Language Processing (NLP) pipeline that performs tokenization, sentence splitting, part-of-speech tagging, and lemmatization using the ILSP suite of NLP tools for Greek (Papageorgiou et al., 2002; Prokopidis et al., 2011). In the next phase, the pre-processing output is given as input to the Semantic Analysis Unit, which performs VA analysis. We employ a rule-based method that comprises of a variety of lexical resources and grammars (sets of linguistic patterns). The VA analyser is a Finite State Transducers (FST) cascade implemented as a JAPE grammar (Cunningham et al., 2000) in the GATE framework. These FSTs process annotation streams utilizing regular expressions to create generalized rules. Moreover, they are ordered in a cascade, so that the output of an FST is given as input to the next transducer. In a first phase the VA analyser detects candidate VAMs and candidate targets based on the respective lexical resources; if a token is recognized as a lexicon entry then it was annotated with the respective metadata (lexicon labels). In particular, the VA analyser comprises of the following lexical resources:

- **TG_lexicon**: contains possible lexicalizations of the TGs (e.g. "μουσουλμάνος" for "Muslims"). Each term is assigned a respective TG_id label.

- **TGVA_lexicon**: contains possible lexicalizations of the TGs that express at the same time VA e.g. racial slurs, derogatory morphological variations of nationality adjectives (e.g. Πακιστανά, Αλβαναριό). Each term is assigned a respective TG_id label, and a semantic label indicating the VAM type it belongs to.

- **VAM1_lexicon**: contains a customized version of EvalLex (Pontiki et al., 2013; Pontiki and Papageorgiou, 2015), an Appraisal Theory (Martin and White, 2005) grounded Lexicon for Evaluative Language that was manually compiled for the Greek language. Each term is assigned a label according to its category (i.e. adjective (JJ), adverb (RB), noun (NN), or verb (VB)) and its sentiment polarity (i.e. negative (n), positive (p), or both (b)). In addition, the terms were further classified as follows based on the strength degree of their evaluative meaning (EM) and prior polarity (PP): 1) Strong EM with a strong (p/n) PP e.g. *"υπεροπτικός" ("arrogant")* [JJ1n]. 2) Weak EM with a strong (p/n) PP e.g. *"ώριμος" ("mature")* [JJ2p]. 3) Strong or weak EM with a weak (p/n/b) or no PP e.g. *"μικρός" ("small")* [JJ3b]. For the needs of the current project each lexicon entry was assigned also a semantic label indicating the VAM type it belongs to.

- **VAM2_lexicon**: contains terms used to express verbal aggression of type 2. Each term was assigned a label according to its category (i.e. adjective (JJ), adverb (RB), noun (NN), or verb (VB)) and the VAM type.

In a subsequent step, the grammars determine which spotted candidate VAMs and targets are correct. The grammars are the implementation of multi-phase algorithms where the output of each phase is input for the next one. Each phase comprises several modules that contain a variety of contextual lexico-syntactic patterns. The patterns are templates that generate rules in the context around the candidate verbal attacks and targets.

For each identified VAM, the method returns the type and the linguistic evidence of the attack as well as of the object of the attack (TG). The output is recorded in the Knowledge Database described below in section 2.5 and is, then, used for statistical analysis and visualizations (section 2.6).

2.5 Population of the Knowledge Database

The Twitter collections described in section 2.1 were automatically processed through the Data Analytics pipeline for VA analysis described in the previous section (2.4). For each processed Tweet the Knowledge Database was populated with two types of metadata following the structure described below:

- Annotations derived by the VA analysis (see above 2.4)
    - **TG_id**: string variable
    - **TG_evidence**: string variable
    - **VAM_type**: string variable
    - **VA_evidence**: string variable

- Twitter metadata
    - Tweet timestamp
        - **Year**: numeric variable
        - **Month**: numeric variable
        - **Day**: numeric variable
    - **User_id**: numeric variable (The Twitter ID of the user that texted the Tweet)
    - **Text**: The actual Tweet message.

The Tweet timestamp was split in three separated fields (Day, Month, Year) instead of one (Day/Month/Year) in order to be able to produce more fine-grained visualizations like timelines and thus, to better monitor the evolution of VA in time.

The information about the User ID can help to identify highly aggressive users as well as to be exploited for social network analysis in order to spot specific communities that promote xenophobic attitudes.

A snapshot of the database is provided below in Figure 6, whilst the total number of the VAMs per TG is presented in Table 2. The database is available through the CLARIN-EL infrastructure, the Greek part of the CLARIN European Research infrastructure through which researchers have access to digital language data. (http://www.clarin.gr/).

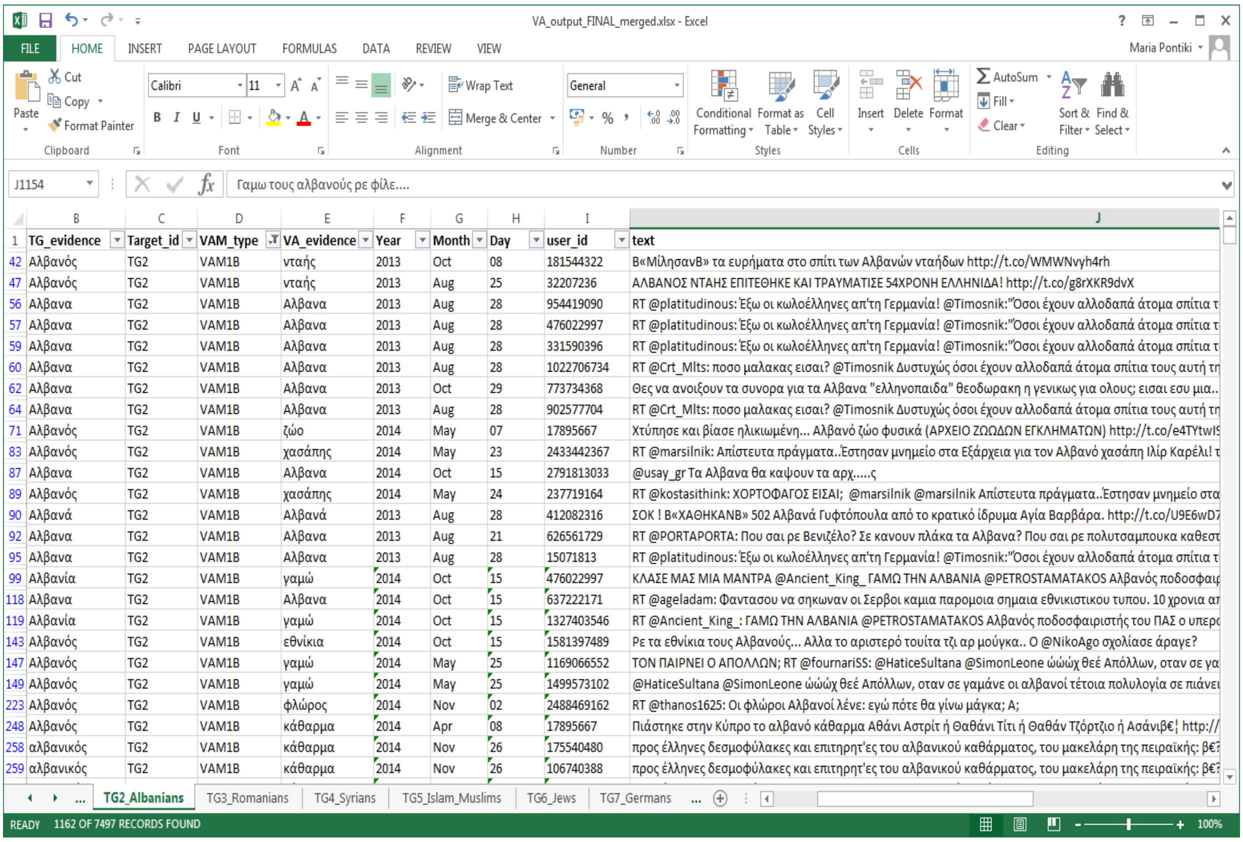

Figure 6: Snapshot of the Knowledge Database

Target Group	Number of VA messages
TG1: Pakistani	**1681**
TG2: Albanians	**7497**
TG3: Romanians	**717**
TG4: Syrians	**633**
TG5: Muslims	**7173**
TG6: Jews	**4050**
TG7: Germans	**10900**
TG8: Roma	**883**
TG9: Immigrants	**11268**
TG0: Refugees	**3011**

Table 2: Amount of VAMs per TG

Finally, the output was visualized in various ways (see below section 2.6) giving a better understanding of the data and the results of the VA analysis.

2.6 Visualization

The content of the Knowledge Database was visualized in various ways in order to make the VA results explorable, comprehensible and thus more easily interpretable. The different types of information types that were extracted, allow for many different associations and graphs for both quantitative and qualitative analysis. In particular, the generated visualizations include graphs, pie charts, timelines, and word clouds. Some examples of the generated visualizations that were used for addressing the research questions of the project are presented below.

- **Graphs** that display the VA analysis results per year and per TG (e.g. Figure 7). Such graphs provide an overview of the most and the least attacked TGs and can help to monitor xenophobia in time (peak points, discontinuities etc.).

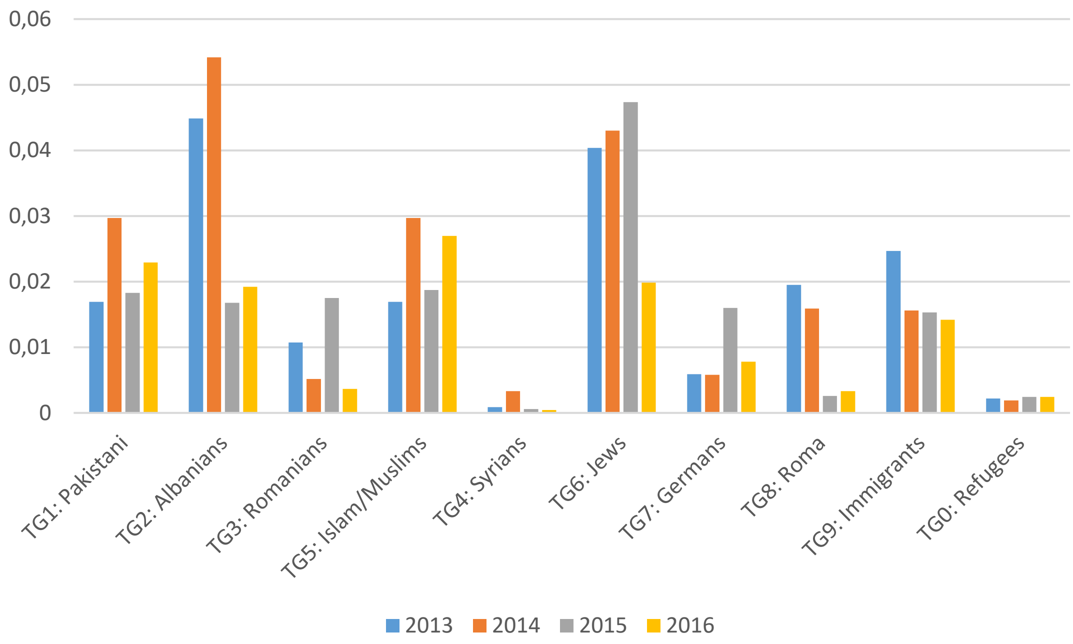

Figure 7: VA rate (VAMs/Tweets) per TG

- **Pie charts** that present the distribution of the different VAM types per TG (e.g. Figures 8 and 9). Such charts can help to explore whether different types of VA can be associated with different TGs, in other words to explore if "foreigners" can be framed based on specific VAM types.

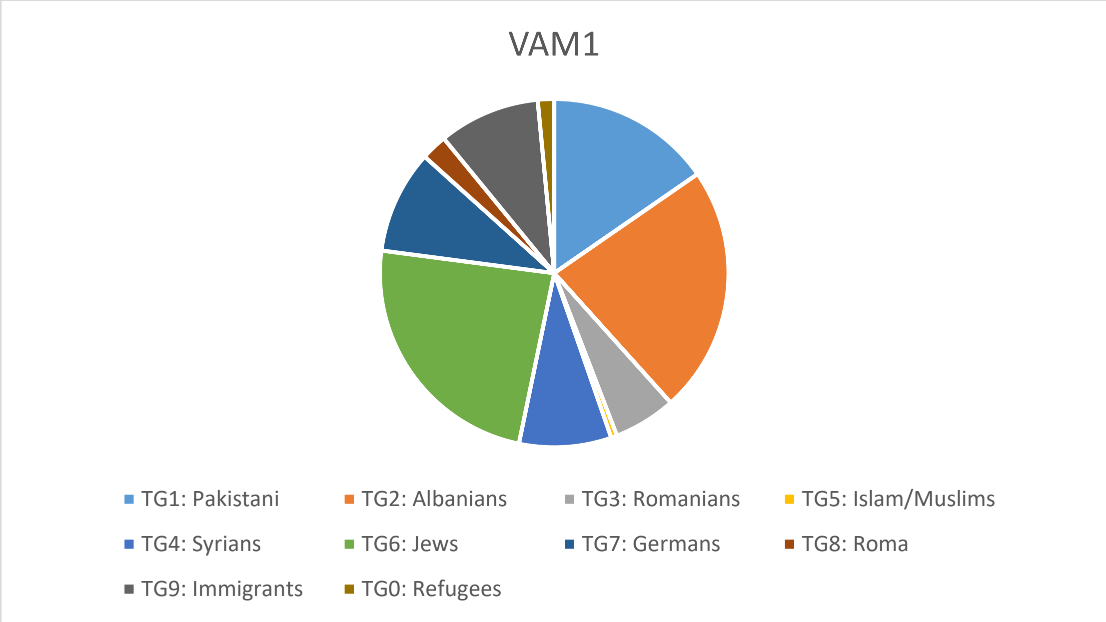

Figure 8: VAM1 rate per TG

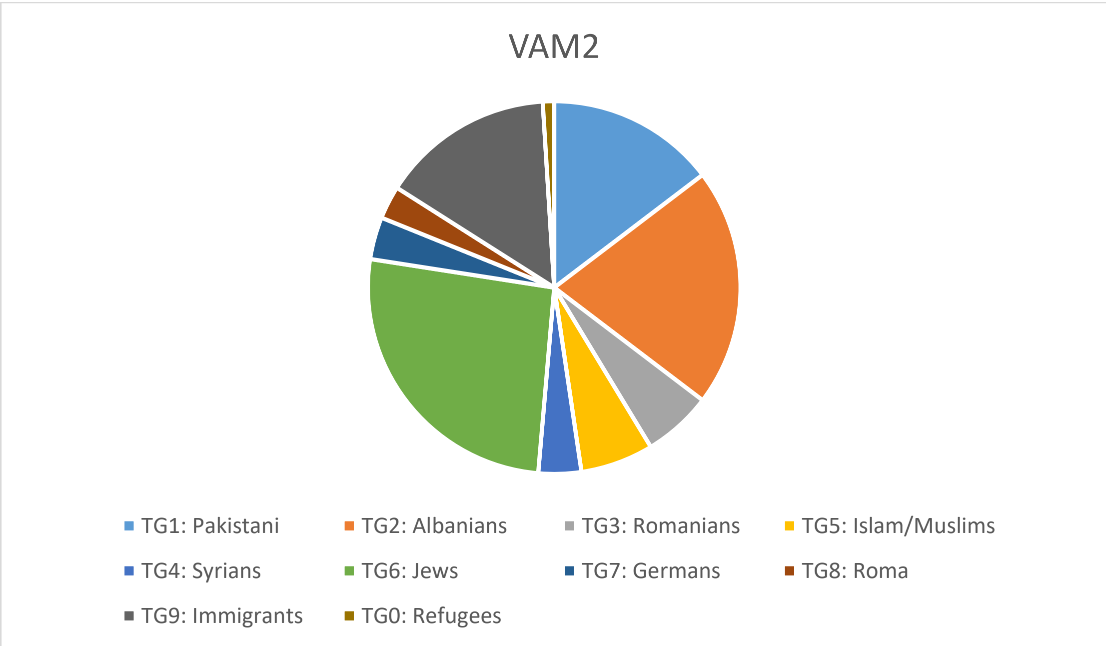

Figure 9: VAM2 rate per TG

- **Word Clouds** that display the unique aggressive terms captured per TG. Clouds of this type make the results understandable and easily usable for the human eye. They are very useful since they can provide access to the different attributes/aspects that are being attacked in each case and, thus, reveal dominant stereotypes per TG. For example, as illustrated in Figure 10, Islam is perceived as a disease (e.g. “γάγγραινα” [“gangrene”]) and is verbally attacked using a variety of unique terms that indicate irrationalism/inferiority (“σκοταδισμός” [“obscurantism”]), sexist behavior (e.g. “μισογυνιστικός” [“woman-hating, misogynistic”]) and fanaticism (e.g. “ισλαμοφασίστες” [“islamofascists”]). On the other hand, the word cloud for Pakistani (Figure 11) contains less unique terms most of which are derogatory morphological variations of the nationality adjective “Πακιστανός” implying inferiority.

Figure 10: Word Cloud of unique aggressive terms for the TG “Muslims/Islam”

Figure 11: Word Cloud of unique aggressive terms for the TG "Pakistani"

3. Conclusions

The current project examines the phenomenon of xenophobia in Greece during the economic crisis. An essential task towards this end is to harvest, exploit and incorporate knowledge from Social Media channels focusing on opinionated user-generated content as a key information source for capturing and understanding xenophobic attitudes. In our approach xenophobia is not merely an attitude, but a form of practice which is "rooted in the symbolic violence of everyday life". To this end, we analyzed Tweets collected during the period 2013-2016 in order to study the formulation of VA in relation to specific TGs of interest (Jews, Muslims, Albanians, etc.) defined based on a number of criteria (e.g. population of the specific ethnic groups in Greece, dominant prejudices in Greece about the specific groups). The automatic detection of verbal attacks helps to measure and monitor xenophobia as a violent practice in Greece over time.

This report provided an overview of the methodology followed for building and populating the knowledge database with the output of the automatic VA analysis of Twitter user generated content. The knowledge database helps to study the formulation of VA in relation to specific TGs, to measure and monitor different aspects of VA in time and provides insights (e.g. aggressive terms, emerging stereotypes) for the research questions the particular project aims to address (Deliverables 3.1 and 4.1). In addition, given the high correlation between verbal and physical aggression (Berkowitz, 1993; Hamilton and Hample, 2011; Laineste, 2012) -in that verbal aggression may escalate to physical violence- the proposed method can provide valuable insights also to policy makers.

References

Allan, K., and Burridge, K. (2006). “Forbidden Words: Taboo and the Censoring of Language”. Cambridge: Cambridge University Press, 2006.

Austin, J. L. (1962). “How to do things with words”. Cambridge (Mass.)

Berkowitz, L. (1993). “Aggression: Its causes, consequences, and control”. New York, NY: McGraw-Hill.

Crystal, D. (1995). “The Cambridge Encyclopedia of the English language”.

Cunningham H., Maynard D. and Tablan V. (2000)/. “JAPE: a Java annotation patterns engine”. Research Memorandum CS–00–10, Department of Computer Science, University of Sheffield.

Graumann, C. (1998). “Verbal discrimination: A neglected chapter in the social psychology of aggression”, *Journal for the Theory of Social Behaviour*, 1998, 28.1, 41-61 (46).

Hamilton, M., and Hample, D. (2011). “Testing Hierarchical Models of Argumentativeness and Verbal Aggressiveness”, Communication Methods and Measures, 2011, 5(3), 250-273.

Infante, D., and Wigley, C. (1986). “Verbal aggressiveness: An interpersonal model and measure”, Communication Monographs, 1986, 53:1, 61-69 (61)

Infante, D. (1987). “Aggressiveness”, in J.C. McCroskey and J.A. Daly (eds.), Personality and interpersonal communication, Sage, Newbury Park 1987, 158.

Infante, D.A., Sabourin, T.C., Rudd, J.E., & Shannon, E. A. (1990). "Verbal Aggression in Violent and Nonviolent Marital Disputes." *Communication Quarterly, 38*(4), p361-371

Kiesler, S., Siegel, J., and McGuire, T. (1984) “Social psychological aspects of computer-mediated communication”. *American Psychologist,* 39 (10): 1123-1134.

Kinney, T. (1994). “An inductively derived typology of verbal aggression and its association to distress”, *Human Communication Research*, 1994, 21:2, 183-222.

Laineste, L. (2012). “Verbal expressions of aggressiveness in Estonian Internet”, Estonia and Poland. Creativity and tradition in cultural communication, 2012, 1: 205-220.

Martin, J.R., and White, P.R.R. (2005). “The Language of Evaluation, Appraisal in English”. Palgrave Macmillan, London & NY.

Mercury, R. E. (1995) “Swearing: A "Bad" Part of Language; a Good Part of Language Learning.

Millard, W. B. (1997). “I flamed Freud: A case study in teletextual incendiarism.”p. 145-159.

Papageorgiou, H., Prokopidis, P., Demiros, I., Giouli, V., Konstantinidis , A., and Piperidis , S. (2002) . “Multi–level XML–based Corpus Annotation.”. In Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC 2002). Las Palmas, Spain.

Papanikolaou, K., Papageorgiou, H., Papasarantopoulos, N., Stathopoulou, T., Papastefanatos, G. (2016). “Just the Facts” with PALOMAR: Detecting Protest Events in Media Outlets and Twitter. In International AAAI Conference on Web and Social Media. North America.

Pontiki, M., Angelou, Z., and Papageorgiou, H. (2013). “Sentiment Analysis: Building Bilingual Lexical Resources”. In Proceeding of the 11th International Conference on Greek Linguistics, September 26-29, Rhodes Island, Greece.

Pontiki, M., and Papageorgiou, H. (2015). *Opinion Mining and Target Extraction in Greek Review Texts. (to appear in)* Proceeding of the 12th International Conference on Greek Linguistics, September 16-19, Berlin, Germany.

Prokopidis, P., Georgantopoulos, B., and Papageorgiou, H. (2011). “A suite of NLP tools for Greek”. In The 10th International Conference of Greek Linguistics. Komotini, Greece.

Reynolds, V., and. Vine, I. (1987). The Socio-Biology of Ethnocentrism: Evolutionary Dimensions of Xenophobia, Discrimination, Racism and Nationalism, Croom Helm, London 1987, 28.

Tereszkiewicz, A. (2012). Do Poles flame? Aggressiveness on Polish discussion groups and social networking sites Jan 2013 · The Journal of Genetic Psychology

van der Veer, K. et al. (2013). “Psychometrically and qualitatively validating a cross-national cumulative measure of fear-based xenophobia”, Quality & Quantity 47.3 (2013): 1429-1444.

Verkhovsky, A. (2006). Sova center report on hate speech in Russian Internet.