Graphic Representations of Spoken Interactions from Journalistic Data: Persuasion and Negotiations

Christina Alexandris[1,2(✉)], Vasilios Floros[1,2], and Dimitrios Mourouzidis[1,2]

[1] National and Kapodistrian University of Athens, Athens, Greece
calexandris@gs.uoa.gr
[2] European Communication Institute (ECI), Danube University Krems and National Technical University of Athens, Athens, Greece

Abstract. Generated graphic representations for interactions involving persuasion and negotiations are intended to assist evaluation, training and decision-making processes and for the construction of respective models. As described in previous research, discourse and dialog structure are evaluated by the y level value around which the graphic representation is developed. Special emphasis is placed on emotion used as a tool for persuasion with the respective expressions, pragmatic elements and the depiction of information not uttered and their subsequent use in the collection of empirical and statistical data.

Keywords: Spoken journalistic texts · Spoken interaction · Persuasion · Negotiations · Graphic representations · Cognitive bias

1 Registration of Spoken Interaction: Previous Research

With the increase in the variety and complexity of spoken Human Computer Interaction (HCI) (and Human Robot Interaction - HRI) applications, the correct perception and evaluation of information not uttered is an essential requirement in systems with emotion recognition, virtual negotiation, psychological support or decision-making. Pragmatic features in spoken interaction and information conveyed but not uttered by Speakers can pose challenges to applications processing spoken texts that are not domain-specific, as in the case of spoken political and journalistic texts, including cases where the elements of persuasion and negotiations are involved.

Although usually underrepresented both in linguistic data for translational and analysis purposes and in Natural Language Processing (NLP) applications, spoken political and journalistic texts may be considered to be a remarkable source of empirical data both for human behavior and for linguistic phenomena, especially for spoken language. However, these text types are often linked to challenges for their evaluation, processing and translation, not only due to their characteristic richness in socio-linguistic and socio-cultural elements and to discussions and interactions beyond a defined agenda, but also in regard to the possibility of different types of targeted audiences - including non-native speakers and the international community [1]. Additionally, in spoken

political and journalistic texts there is also the possibility of essential information, presented either in a subtle form or in an indirect way, being often undetected, especially by the international public. In this case, spoken political and journalistic texts also contain information that is not uttered but can be derived from the overall behavior of speakers and participants in a discussion or interview. These characteristics, including the feature of spontaneous turn-taking [31, 39] in many spoken political and journalistic texts, are linked to the implementation of strategies concerning the analysis and processing of discourse structure and rhetorical relations (in addition to previous research) [10, 22, 35, 41].

In our previous research [2, 6, 23], a processing and evaluation framework was proposed for the generation of graphic representations and tags corresponding to values and benchmarks depicting the degree of information not uttered and non-neutral elements in Speaker behavior in spoken text segments. The implemented processing and evaluation framework allows the graphic representation to be presented in conjunction with the parallel depiction of speech signals and transcribed texts. Specifically, the alignment of the generated graphic representation with the respective segments of the spoken text enables a possible integration in existing transcription tools.

In particular, strategies typically employed in the construction of most Spoken Dialog Systems, such as keyword processing in the form of topic detection [13, 19, 24, 25] (from which approaches involving neural networks are developed [38]), were adapted in the functions of the designed and constructed interactive annotation tool [2, 6, 23], designed to operate with most commercial transcription tools. The output provides the User-Journalist with (a) the tracked indications of the topics handled in the interview or discussion and (b) the graphic pattern of the discourse structure of the interview or discussion. The output (a) and (b) also included functions and respective values reflecting the degree in which the speakers-participants address or avoid the topics in the dialog structure ("RELEVANCE" Module) as well as the degree of tension in their interaction ("TENSION" Module).

The implemented "RELEVANCE" Module [23], intended for the evaluation of short speech segments, generates a visual representation from the user's interaction, tracking the corresponding sequence of topics (topic-keywords) chosen by the user and the perceived relations between them in the dialog flow. The generated visual representations depict topics avoided, introduced or repeatedly referred to by each Speaker-Participant, and in specific types of cases may indicate the existence of additional, "hidden"[23] Illocutionary Acts [9, 14, 15, 32] other than "Obtaining Information Asked" or "Providing Information Asked" in a discussion or interview.

Thus, the evaluation of Speaker-Participant behavior targets to by-pass Cognitive Bias, specifically, Confidence Bias [18] of the user-evaluator, especially if multiple users-evaluators may produce different forms of generated visual representations for the same conversation and interaction. The generated visual representations for the same conversation and interaction may be compared to each other and be integrated in a database currently under development. In this case, chosen relations between topics may describe Lexical Bias [36] and may differ according to political, socio-cultural and linguistic characteristics of the user-evaluator, especially if international users are concerned [21, 26, 27, 40] due to lack of world knowledge of the language community

involved [7, 16, 37]. In the "RELEVANCE" Module [23], a high frequency of Repetitions (value 1), Generalizations (value 3) and Topic Switches (value -1) in comparison to the duration of the spoken interaction is connected to the "(Topic) Relevance" benchmarks with a value of "Relevance (X)" [3, 5] (Fig. 1).
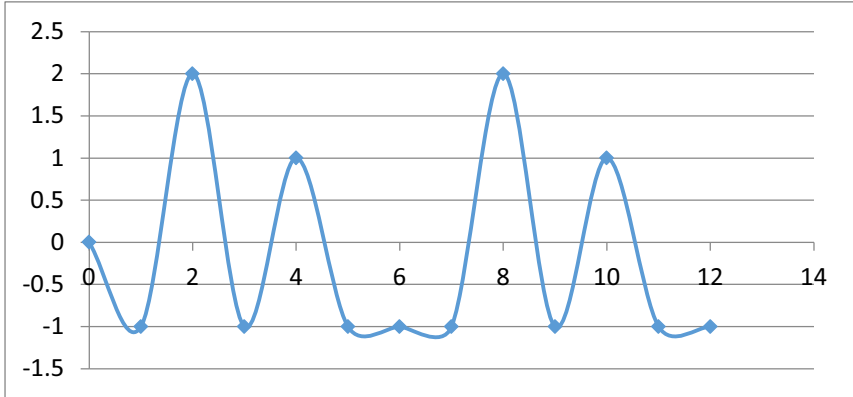

Fig. 1. Generated graphic representation with multiple "Topic Switch" relations (Mourouzidis et al., 2019).

The development of the interactive, user-friendly annotation tool is based on data and observations provided by professional journalists (European Communication Institute (ECI), Program M.A in Quality Journalism and Digital Technologies, Danube University at Krems, Austria, the Athena- Research and Innovation Center in Information, Communication and Knowledge Technologies, Athens, the Institution of Promotion of Journalism Ath.Vas. Botsi, Athens and the National and Technical University of Athens, Greece).

2 Association Relations and (Training) Data for Negotiation Models

However, in the above-presented previous research, the "Association" relation is not included in the evaluations concerned. Furthermore, the "Association" relation is of crucial importance in dialogues constituting persuasion and types of negotiation based on persuasion [34], especially if emotion is used as a tool for persuasion [30], establishing a link between persuasion, emotion and language [30]. Emotion as a tool for persuasion may be used in diverse types of negotiation skills, apart from persuasion tactics [12, 30, 34], including "value creating"/ "value claiming" tactics and "defensive" tactics [34].

"Association" relations between words and their related topics are often used to direct the Speaker into addressing the topic of interest and/or to produce the desired answers. In some cases, the "Generalization" may also be used for the same purpose, as a means of introducing a (not directly related) topic of interest via "Generalization".

For negotiation applications, the identification of words and their related topics contributes to strategies targeting to directing the Speaker-Participant to the desired goal and the avoidance of unwanted "Association" types as well as unwanted other types of relations -"Repetitions", "Topic Switch" and "Generalizations" (Fig. 2).
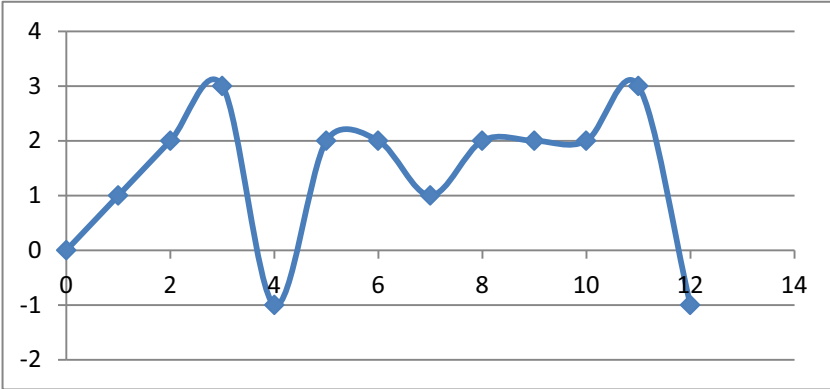

Fig. 2. Generated graphic representation with multiple "Association" relations. (Mourouzidis et al., 2019).

The "Association" relations between words and their related topics contribute to the analysis and development of negotiation procedures. In this case, Cognitive Bias and socio-cultural factors play a crucial role in regard to the perception of the perceived relations-distances between word-topics. For example, the word-topics "Country X" (name withheld) –"defense spending" or "military confrontation" – "chemical weapons" may generate an "Association" (ASOC) or "Topic Switch" (SWITCH) reactions and choices from users, depending on whether they are perceived as related or different topics in the spoken interaction. Diverse reactions may also apply in the case of the "Association" and "Generalization" relations, where "treaties" and "international commitment" may generate "Association" (ASOC) or "Generalization" (GEN) reactions and choices from users: "treaties" is associated with "international commitment" or "treaties" are linked to "international commitment" with a "Generalization" relation.

Differences concerning the perception of the "Association" (ASOC) relations between word-topics are measured in the form of triple tuples as perceived relations-distances between word-topics [3], related to Lexical Bias (Cognitive Bias) concerning semantic perception [36]. Examples of segments in (interactively) generated patterns from user-specific choices between topics are the following, where the distances between topics in the generated patterns are registered as triple tuples (triplets): (military confrontation, chemical weapons, 2) ("Association"), (treaties, international commitment, 3) ("Generalization"). These triplets and the sequences they form may be converted into vectors (or other forms and models), used as training data for creating negotiation models and their variations.
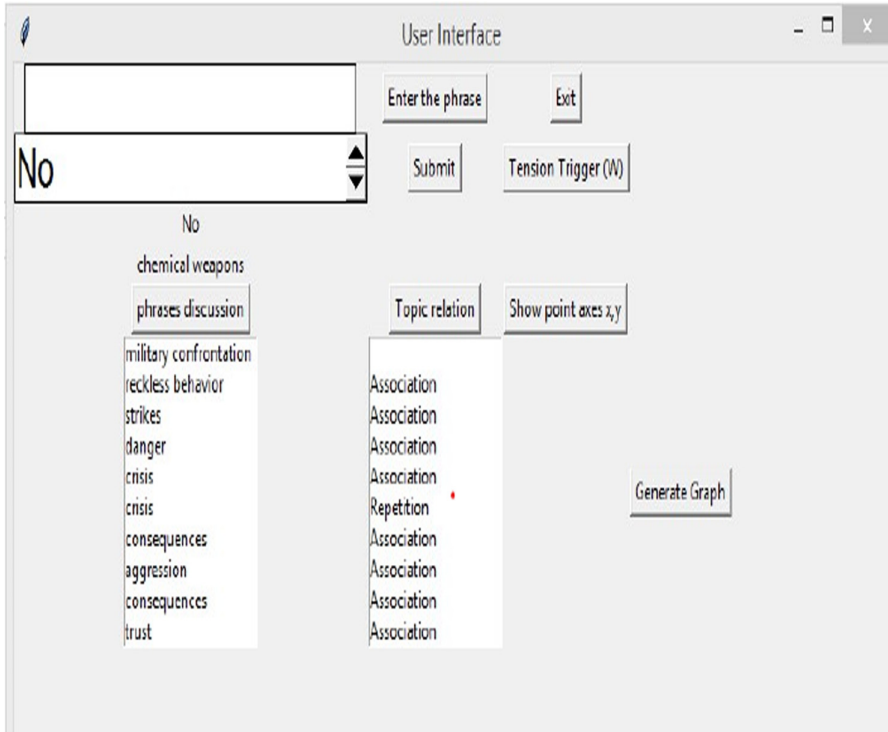
Fig. 3. Interface for generating graphic representation with multiple "Association" relations.

Possible differences in the perceived relations with the Lexical Bias concerned may play an essential role both in the employment of negotiation tactics (based on cross-cultural analysis) and in training applications. The number of registered "Association" relations in the processed wav.file or video file may be used to evaluate persuasion tactics employed in spoken interaction involving negotiations (a) and their possible employment in the construction of training data and negotiation models (b). Since the generated graphic representations are based on perceived relations, they may also be used for evaluating trainees performance (c).

We note that, independently from interactive and user-specific choices, topics may be also pre-defined and/or automatically detected with word relations based on existing (ontological and semantic) databases. However, this commonly used strategy and practice is proposed to be employed in cases where persuasion and negotiation tactics are monitored and checked against a pre-defined model, either as a form to control spoken interaction or as means to evaluate the pre-defined model.

The following examples in Figs. 3 and 4 depict the user interface and the generated graphic representations containing multiple "Association" relations: Chosen word-topics and their relations in dialog segment with two speakers-participants (resulting to a "No" answer): "military confrontation", "reckless behavior", "strikes", "danger", "crisis", "crisis", "consequences", "aggression", "consequences", "trust". (choices may

vary among users, especially in the international public), Data from an actual interview on a world news channel (BBC HardTalk 720- 16–04-2018).
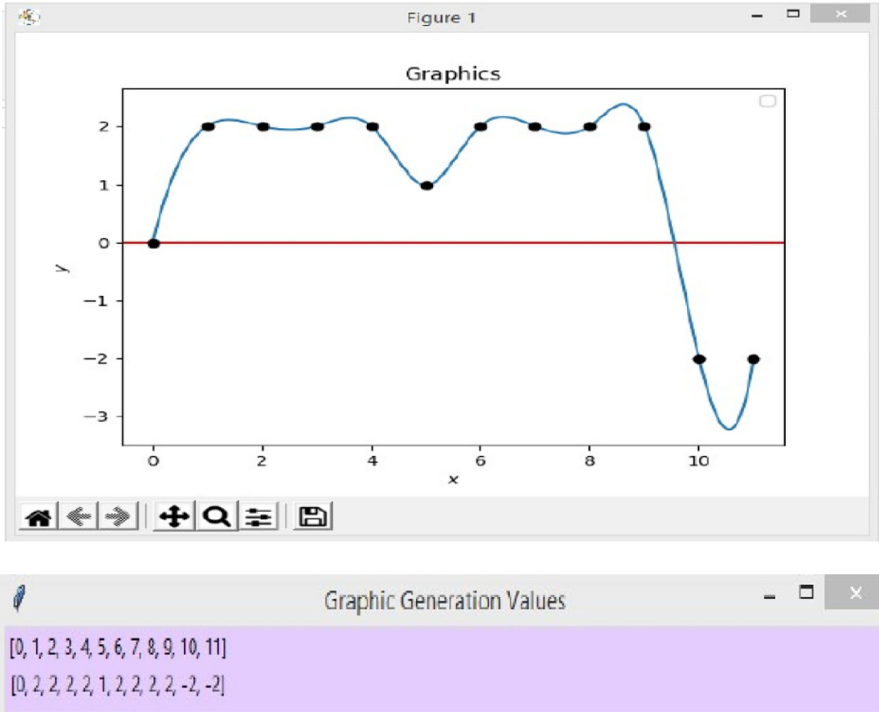

Fig. 4. Generated graphic representation with multiple "Association" relations and respective values (including one "No" Answer (−2) – presented in Sect. 3).

3   Affirmative and Negative Answers in Negotiations

In spoken interaction concerning persuasion and types of negotiation based on persuasion [12, 30, 34], perceived affirmative ("Yes") and negative ("No") answers are integrated in the present framework with the respective "0" (zero) and "−2" values.

Specifically, an affirmative answer is assigned a "0" value, similar to the initial "0" (zero) value starting the entire interactive processing of the wav.file. An example of a generated graphic representation with multiple "Yes" answers is depicted in Fig. 5. In this case, the spoken interaction (concerning persuasion or negotiation based on persuasion) contains multiple positive answers and the respective multiple "0" (zero) values (Fig. 5).

A negative answer is assigned a "−2" value, lower than the "−1" Topic Switch value (Fig. 6). Thus, a negotiation with a sequence of negative answers and several attempts to change a topic or to approach a (seemly) different topic will generate a graphic representation below the "0" (zero) value.

An example of generated graphic representations below the "0" (zero) value depicting spoken interactions (persuasion –negotiations) is shown in Fig. 6. In this case, the spoken interaction contains multiple negative answers and/or multiple attempts to switch to a different topic (Fig. 6).
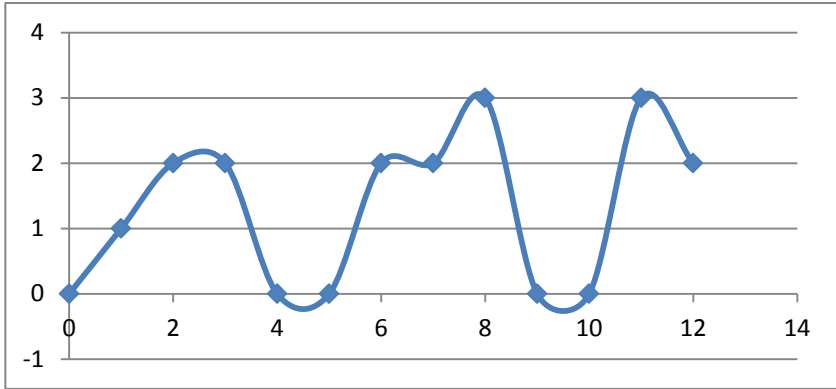

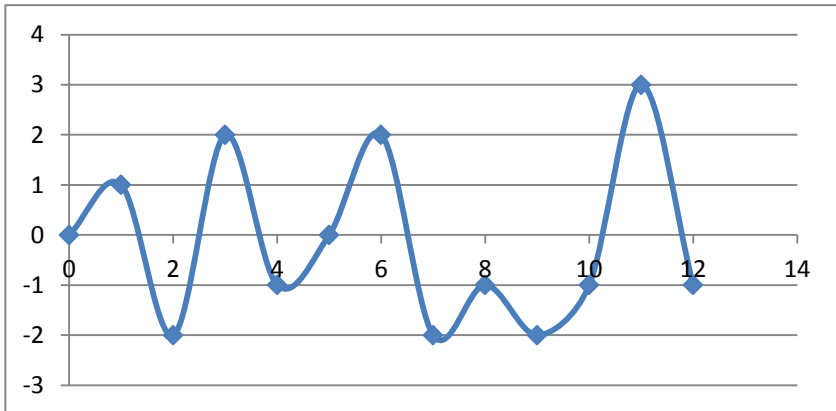**Fig. 5.** Generated graphic representation with multiple "Yes" answers.

Fig. 6. Generated graphic representation with multiple "No" answers (and topic switches).

As in the above-described case of "Association" and "Generalization" relations, for affirmative and negative answers, the distances between topics in the generated patterns are registered and may be be used as training data for creating negotiation models and their variations. However, in the case of affirmative and negative answers, the topic and the respective answer is not registered as a triplet but is registered as a tuple: (stability, 0) ("Affirmative Answer"), (sanctions, −2) ("Negative Answer").

Similarly to the registered "Association" relations, the number of perceived affirmative ("Yes") and negative ("No") answers in the processed wav.file or video file may be

used to evaluate persuasion tactics employed in spoken interaction involving negotiations (a), for the construction of training data and negotiation models (b) or for evaluating a trainees performance (c).

4 Registering Word-Topics and Their Impact in Persuasion and Negotiations

4.1 Word-Topics and Persuasion Tactics

The type of word-topics concerned in the registered "Association" relations and the "Yes" or "No" answers in the processed wav.file or video file may also be used to evaluate persuasion tactics employed in spoken interaction involving negotiations. Word-topics and the registered relations and answers may be linked to positive responses and/or collaborative speaker behavior or negative responses, tension and conflict. Detecting and registering points of tension or other types of behavior and their impact in the dialogue structure facilitates the evaluation of persuasion tactics and types of negotiation based on persuasion [30, 34], especially "value creating"/ "value claiming" tactics and "defensive" tactics [30, 34] and in other cases where a link between persuasion, emotion and language is used [12, 30].

4.2 Word-Topics and Word-Types as Reaction Triggers

For negotiation applications, words and their related topics can be identified as triggers for different types of reactions (positive, collaborative behavior or tension). The words and their related topics may concern the following two types of information: (1) "Association" (or other) relations that are context-specific, connected to current events and state-of-affairs, (2) "Association" (or other) relations that concern words with inherent socio-culturally determined linguistic features and are usually independent from current events and state-of-affairs.

In the second case (2) it is often observed that the semantic equivalent of the same word on one language sometimes may appear more formal or with more "gravity" than in another language, either emphasizing the role of the word in an utterance or being related to word play and subtle suggested information. The presence of such "gravity words" [1, 4] may contribute to the degree of formality or intensity of conveyed information in a spoken utterance. It is observed that these differences between languages in regard to the "gravity" of words are often related to polysemy, where the possible meanings and uses of a word seem to "cast a shadow" over its most commonly used meaning. Similarly to the above-described category, words with an "evocative" element concern their "deeper" meanings related to their use in tradition, in music and in literature and may sometimes be related to emotional impact in discussions and speeches. In contrast to "gravity" words, "evocative" words usually contribute to a descriptive or emotional tone in an utterance [1, 4]. Here, it is noted that, according to Rockledge et al., 2018, "the more extremely positive the word, the greater the probability individuals were to associate that word with persuasion" [30].

In the generated graphic representations, perceived "Gravity" and "Evocative" words are signalized (for example, as "W") in the curve connecting the word-topics. This

signalization indicates the points of "Gravity" and "Evocative" words as "Word-Topic" triggers in respect to the areas of perceived tension or other types of reactions in the processed dialog segment with two (or more) speakers-participants. In Figs. 7 and 8 the perceived "Gravity" and "Evocative" words also constitute word-topics (Figs. 7 and 8).
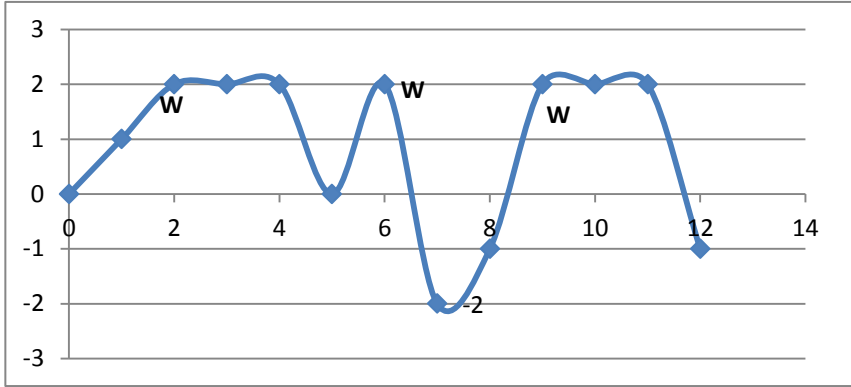

Fig. 7. Generated graphic representation with multiple "Association" relations and Word-Topic triggers ("W").
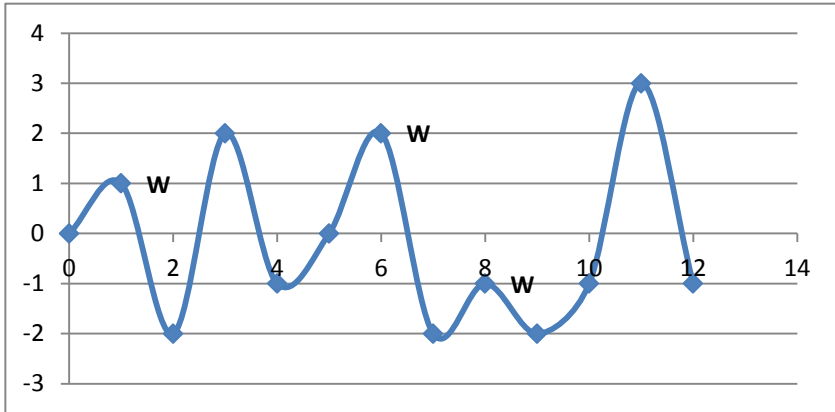

Fig. 8. Generated graphic representation with multiple "No" answers and Word-Topic triggers ("W").

The detected word types may be used as training data for creating negotiation models and their variations, as in the above-described cases. The signalized Word-Topic triggers may be appended as marked values (for example, with "&") in the respective tuples or triple tuples, depending on the context in which they occur: (sanctions, −2, &dignity) ("Negative Answer"), (military confrontation, chemical weapons, 2, &justice) ("Association"). If the Word-Topic triggers constitute topics, they are repeated in the tuple or triple tuple, where they receive the respective mark: (country, people, 2, &people) ("Association").

Signalized "Gravity" and "Evocative" words can be identified either from databases constructed from collected empirical data or from existing resources such as Wordnets.

In spoken utterances "Gravity" words and especially "Evocative" words are observed to often have their prosodic and even their phonetic-phonological features intensified [1, 4]. The commonly occurring observed connection to intensified prosodic phonetic-phonological features constitutes an additional pointer to detecting and signalizing "Gravity" and "Evocative" words [1, 4].

4.3 Word-Topics as Tension Triggers

Previous research depicted points of tension in two-party discussions and interviews containing longer speech segments. These points are detected and signalized by the implemented "TENSION" Module in the form of graphic representations [2], enabling the evaluation of the behavior of speakers-participants.

In spoken interaction concerning persuasion and types of negotiation based on persuasion, detected points of tension in the generated graphic representations enable the registration of word-topics and sequences of word-topics preceding tension and the registration of word-topics and sequences of word-topics following tension. The evaluation of such data contributes both to the construction and training of models for the avoidance of tension (i) and for the purposeful creation of tension (ii).

Multiple points of tension (referred to as "hot spots") [2] indicate a more argumentative than a collaborative interaction, even if speakers-participants display a calm and composed behavior. Points of possible tension and/or conflict between speakers-participants ("hot-spots") are signalized in generated graphic representations of registered negotiations (or other type of spoken interaction concerning persuasion), with special emphasis on words and topics triggering tension and non-collaborative speaker-participant behavior.

As presented in previous research [2], a point of tension or "hot spot" consists of the pair of utterances of both speakers, namely a question-answer pair or a statement-response pair or any other type of relation between speaker turns. In longer utterances, a defined word count and/or sentence length from the first words/segment of the second speaker's (Speaker 2) and from the words/segment of the first speaker's (Speaker 1) the utterance are processed [2, 11]. The automatically signalized "hot spots" (and the complete utterances consisting of both speaker turns) are extracted to a separate template for further processing. For a segment of speaker turns to be automatically identified as a "hot spot", a set of (at least two of the proposed three (3) conditions must apply [2] to one or to both of the speaker's utterances. The three (3) conditions are directly or indirectly related to flouting of Maxims of the Gricean Cooperative Principle [14, 15] (additional, modifying features (1), reference to the interaction itself and to its participants with negation (2) and (3) prosodic emphasis and/or exclamations). With the exception of prosodic emphasis, these conditions concern features detectable with a POS Tagger (for example, the Stanford POS Tagger, http://nlp.stanford.edu/software/tagger.shtml) or they may constitute a small set of entries in a specially created lexicon or may be retrieved from existing databases or Wordnets. The "hot spots" are connected to the "Tension" benchmark with a value of "Y" or "Tension (Y)" [2] and the "Collaboration"

benchmark with a value of "Z" or "Collaboration (Z)", described in previous research [2, 3].

In the generated graphic representations, word-topics as tension triggers are signalized (for example, as "W") in the curve connecting the word-topics (Fig. 9). This signalization indicates the points of word-topics as tension triggers in respect to the areas of perceived tension in the processed dialog segment with two (or more) speakers-participants. The detected word types may be used as training data for creating negotiation models and their variations, as in the above-described cases.
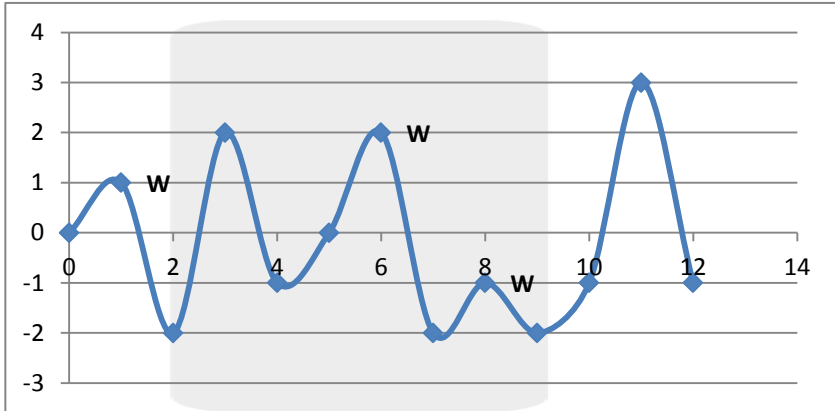

Fig. 9. Generated graphic representation with multiple "No" answers and Word-Topic triggers ("W") and Tension (shaded area between topics) in generated graphic representation and "tension trigger" ("W").

4.4 Tension Triggers and Paralinguistic Information

Furthermore, in previous research [2] "hot spots" signalizing tension may include an interactive annotation of paralinguistic features with the corresponding tags. Words classified as "tension triggers" may, in some cases, be easily detected with the aid of registered and annotated paralinguistic features, where the paralinguistic element may complement or intensify the information content of the word related to perceived tension in the spoken interaction. In some instances, the paralinguistic element may contradict the information content of the "tension trigger", for example, a smile when a word of negative content is uttered. In this case, the speaker's behavior may be related to irony or a less intense negative emotion such as annoyance or contempt. With paralinguistic features concerning information that is not uttered, the Gricean Cooperative Principle is violated if the information conveyed is perceived as not complete (Violation of Quantity or Manner) or even contradicted by paralinguistic features (Violation of Quality).

Depending on the type of specifications used, for paralinguistic features depicting contradictory information to the information content of the spoken utterance, the additional signalization of "!" is proposed, for example, "[! facial-expr: eye-roll]" and "[! gesture: clenched-fist]".

According to the type of linguistic and paralinguistic features signalized, features of more subtle emotions can be detected. Less intense emotions are classified in the middle and outer zones of the Plutchik Wheel of Emotions [28] and are usually too subtle to be easily extracted by sensor and/or speech signal data. In this case, linguistic information with or without a link to paralinguistic features demonstrates a more reliable source of a speaker's attitude, behavior and intentions, especially for subtle negative reactions in the Plutchik Wheel of Emotions, namely "Apprehension", "Annoyance", "Disapproval", "Contempt", "Aggressiveness" [28]. These subtle emotions are of importance in spoken interactions involving persuasion and negotiations.

Data from the interactive annotation of paralinguistic features may also be integrated into models and training data, however, further research is necessary for the respective approaches and strategies.

5   Conclusions and Further Research: Insights for Sentiment Analysis Applications

The presented generated graphic representations for interactions involving persuasion and negotiations are intended to assist evaluation, training and decision-making processes and for the construction of respective models. In particular, the graphic representations generated from the processed wav.file or video files may be used to evaluate persuasion tactics employed in spoken interaction involving negotiations (a), their possible employment in the construction of training data and negotiation models (b) and for evaluating a trainee's performance (c).

New insights are expected to be obtained by the further analysis and research in the persuasion-negotiation data processed. Further research is also expected to contribute to the overall improvement of the graphical user interface (GUI), as one of the basic envisioned upgrades of the application.

The presented generated graphic representations enable the visibility of information not uttered, in particular, tension and the overall behavior of speakers-participants. The visibility of all information content, including information not uttered, contributes to the collection and compilation of empirical and statistical data for research and/or for the development of HCI- HRI Sentiment Analysis and Opinion Mining applications, as (initial) training and test sets or for Speaker (User) behavior and expectations. This is of particular interest in cases where an international public is concerned and where a variety of linguistic and socio-cultural factors is included.

Information that is not uttered is problematic in Data Mining and Sentiment Analysis-Opinion Mining applications, since they mostly rely on word groups, word sequences and/or sentiment lexica [20], including recent approaches with the use of neural networks [8, 17, 33], especially if Sentiment Analysis from videos (text, audio and video) is concerned. In this case, even if context dependent multimodal utterance features are extracted, as proposed in recent research [29], the semantic content of a spoken utterance may be either complemented or contradicted by a gesture, facial expression or movement. The words and word-topics triggering non-collaborative behavior and tension ("hot spots") and the content of the extracted segments where tension is detected provide

insights for word types and the reaction of speakers, as well as insights of Opinion Mining and Sentiment Analysis.

The above-observed additional dimensions of words in spoken interaction, especially in political and journalistic texts, may also contribute to the enrichment of "Bag-of-Words" approaches in Sentiment Analysis and their subsequent integration in training data for statistical models and neural networks.

References

1. Alexandris, C.: Issues in Multilingual Information Processing of Spoken Political and Journalistic Texts in the Media and Broadcast News, Cambridge Scholars, Newcastle upon Tyne, UK (2020)
2. Alexandris, C., Mourouzidis, D., Floros, V.: Generating graphic representations of spoken interactions revisited: the tension factor and information not uttered in journalistic data. In: Kurosu, M. (ed.) HCII 2020. LNCS, vol. 12181, pp. 523–537. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-49059-1_39
3. Alexandris, C.: Evaluating cognitive bias in two-party and multi-party spoken interactions. In: Proceedings from the AAAI Spring Symposium, Stanford University (2019)
4. Alexandris, C.: Visualizing Pragmatic Features in Spoken Interaction: Intentions, Behavior and Evaluation. In: Proceedings of the 1st International Conference on Linguistics Research on the Era of Artificial Intelligence – LREAI, Dalian, October 25–27, 2019, Dalian Maritime University (2019)
5. Alexandris, C.: Measuring cognitive bias in spoken interaction and conversation: generating visual representations. In: Beyond Machine Intelligence: Understanding Cognitive Bias and Humanity for Well-Being AI Papers from the AAAI Spring Symposium Stanford University, Technical Report SS-18-03, pp. 204-206 AAAI Press Palo Alto, CA (2018)
6. Alexandris, C., Nottas, M., Cambourakis, G.: Interactive evaluation of pragmatic features in spoken journalistic texts. In: Kurosu, M. (ed.) HCI 2015. LNCS, vol. 9171, pp. 259–268. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-21006-3_26
7. Alexandris, C.: English, german and the international "semi-professional" translator: a morphological approach to implied connotative features. J. Lang. Transl. Sejong Univ. Korea **11**(2), 7–46 (2010)
8. Arockiaraj, C.M.: Applications of neural networks in data mining. Int. J. Eng. Sci. **3**(1), 8–11 (2013)
9. Austin, J.L.: How to Do Things with Words, 2nd edn. University Press, Oxford Paperbacks, Oxford (1962).(Urmson, J.O., Sbisà, M. (eds.) 1976)
10. Carlson, L., Marcu, D., Okurowski, M. E.: Building a discourse-tagged corpus in the framework of rhetorical structure theory. In: Proceedings of the 2nd SIGDIAL Workshop on Discourse and Dialogue, Eurospeech 2001, Denmark, September 2001 (2001)
11. Cutts, M.: Oxford Guide to Plain English, 4th edn. Oxford University Press, Oxford, UK (2013)
12. Evans, N.J., Park, D.: Rethinking the persuasion knowledge model: schematic antecedents and associative outcomes of persuasion knowledge activation for covert advertising. J. Curr. Issues Res. Advert. **36**(2), 157–176 (2015). https://doi.org/10.1080/10641734.2015.1023873
13. Floros, V., Mourouzidis, D.: Multiple Task Management in a Dialog System for Call Centers. Master's thesis, Department of Informatics and Telecommunications, National University of Athens, Greece (2016)
14. Grice, H.P.: Studies in the Way of Words. Harvard University Press, Cambridge, MA (1989)

15. Grice, H.P.: Logic and conversation. In: Cole, P., Morgan, J. (eds.) Syntax and Semantics, vol. 3, pp. 41–58. Academic Press, New York (1975)
16. Hatim, B.: Communication Across Cultures: Translation Theory and Contrastive Text Linguistics. University of Exeter Press, Exeter, UK (1997)
17. Hedderich, M.A., Klakow, D.: Training a neural network in a low-resource setting on automatically annotated noisy data. In: Proceedings of the Workshop on Deep Learning Approaches for Low-Resource NLP, Melbourne, Australia, pp. 12–18. Association for Computational Linguistics-ACL (2018)
18. Hilbert, M.: Toward a synthesis of cognitive biases: how noisy information processing can bias human decision making. Psychol. Bull. **138**(2), 211–237 (2012)
19. Lewis, J.R.: Introduction to Practical Speech User Interface Design for Interactive Voice Response Applications, IBM Software Group, USA, Tutorial T09 presented at HCI 2009 San Diego. CA, USA (2009)
20. Liu, B.: Sentiment Analysis and Opinion Mining. Morgan & Claypool, San Rafael, CA (2012)
21. Ma, J.: A comparative analysis of the ambiguity resolution of two English-Chinese MT approaches: RBMT and SMT. Dalian Univ. Technol. J. **31**(3), 114–119 (2010)
22. Marcu, D.: Discourse trees are good indicators of importance in text. In: Mani, I., Maybury, M. (eds.) Advances in Automatic Text Summarization, pp. 123–136. The MIT Press, Cambridge, MA (1999)
23. Mourouzidis, D., Floros, V., Alexandris, C.: Generating graphic representations of spoken interactions from journalistic data. In: Kurosu, M. (ed.) HCII 2019. LNCS, vol. 11566, pp. 559–570. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-22646-6_42
24. Nass, C., Brave, S.: Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship. The MIT Press, Cambridge, MA (2005)
25. Nottas, M., Alexandris, C., Tsopanoglou, A., Bakamidis, S.: A hybrid approach to dialog input in the citzenshield dialog system for consumer complaints. In: Proceedings of HCI 2007, Beijing, People's Republic of China (2007)
26. Paltridge, B.: Discourse Analysis: An Introduction. Bloomsbury Publishing, London (2012)
27. Pan, Y.: Politeness in Chinese Face-to-Face Interaction. Advances in Discourse Processes series, vol. 67. Ablex Publishing Corporation, Stamford, CT, USA (2000)
28. Plutchik, R.: A psychoevolutionary theory of emotions. Soc. Sci. Inf. **21**, 529–553 (1982). https://doi.org/10.1177/053901882021004003
29. Poria, S., Cambria, E., Hazarika, D., Mazumder, N., Zadeh, A., Morency, L-P.: Context-dependent sentiment analysis in user-generated videos. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, Vancouver, Canada, July 30-August 4 2017, pp. 873–88. Association for Computational Linguistics - ACL (2017). https://doi.org/10.18653/v1/P17-1081
30. Rocklage, M., Rucker, D., Nordgren, L.: Persuasion, emotion, and language: the intent to persuade transforms language via emotionality. Psychol. Sci. **29**(5), 749–760 (2018). https://doi.org/10.1177/0956797617744797
31. Sacks, H., Schegloff, E.A., Jefferson, G.: A simplest systematics for the organization of turn-taking for conversation. Language **50**, 696–735 (1974)
32. Searle, J.R.: Speech Acts: An Essay in the Philosophy of Language. Cambridge University Press, Cambridge, MA (1969)
33. Shah, K., Kopru, S., Ruvini, J-D.: Neural network based extreme classification and similarity models for product matching. In: Proceedings of NAACL-HLT 2018, New Orleans, Louisiana, June 1-6, 2018, pp. 8–15. Association for Computational Linguistics-ACL (2018)
34. Skonk, K.: 5 Types of Negotiation Skills, Program on Negotiation Daily Blog, Harvard Law School, 14 May 2020. https://www.pon.harvard.edu/daily/negotiation-skills-daily/types-of-negotiation-skills/. Accessed 11 Nov 2020

35. Stede, M., Taboada, M., Das, D.: Annotation Guidelines for Rhetorical Structure Manuscript. University of Potsdam and Simon Fraser University, Potsdam (2017)
36. Trofimova, I.: Observer bias: an interaction of temperament traits with biases in the semantic perception of lexical material. PLoS ONE **9**(1), e85677 (2014). https://doi.org/10.1371/journal.pone.0085677
37. Wardhaugh, R.: An Introduction to Sociolinguistics, 2nd edn. Blackwell, Oxford, UK (1992)
38. Williams, J.D., Asadi, K., Zweig, G.: Hybrid code networks: practical and efficient end-to-end dialog control with supervised and reinforcement learning. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, Vancouver, Canada, July 30-August 4 2017, pp. 665–677. Association for Computational Linguistics (ACL) (2017)
39. Wilson, M., Wilson, T.P.: An oscillator model of the timing of turn taking. Psychon. Bull. Rev. **12**(6), 957–968 (2005)
40. Yu, Z., Yu, Z., Aoyama, H., Ozeki, M., Nakamura, Y.: Capture, recognition, and visualization of human semantic interactions in meetings. In: Proceedings of PerCom, Mannheim, Germany, 2010 (2010)
41. Zeldes, A.: rstWeb - a browser-based annotation interface for rhetorical structure theory and discourse relations. In: Proceedings of NAACL-HLT 2016 System Demonstrations. San Diego, CA, pp. 1–5 (2016). http://aclweb.org/anthology/N/N16/N16-3001.pdf