



# Generating Graphic Representations of Spoken Interactions from Journalistic Data

Dimitrios Mourouzidis<sup>1</sup>, Vasilios Floros<sup>1</sup>,  
and Christina Alexandris<sup>1,2,3(✉)</sup>

<sup>1</sup> National and Kapodistrian University of Athens, Athens, Greece  
mourouzidisd@gmail.com, florosbas2002@yahoo.gr,  
calexandris@gs.uoa.gr

<sup>2</sup> European Communication Institute (ECI),  
Danube University Krems, Krems, Austria

<sup>3</sup> National and Technical University of Athens, Athens, Greece

**Abstract.** Generated distinct types of graphic patterns depicting the discourse structure of spoken journalistic texts - interviews or discussions, contribute to a user-independent evaluation of spoken Human-Human conversation and interaction. Distinct types of the graphic representation also function as visual representations of Cognitive Bias. The distinct types of generated graphic representations and values allow the identification of additional, “hidden,” Illocutionary Acts, beyond the defined framework of the spoken conversation and interaction.

**Keywords:** Spoken journalistic texts · Spoken interaction · Dialog flow · Graphic representations · Cognitive bias

## 1 Introduction

Generating distinct types of graphic patterns depicting the discourse structure of spoken journalistic texts targets to contribute to a user-independent evaluation of spoken Human-Human conversation and interaction. The generated types of graphic patterns describing the discourse structure of an interview or discussion function as visual aids for the evaluation of pragmatic features and the detection of Cognitive Bias. Furthermore, the graphic patterns generated may also enable the identification of additional, “hidden,” Illocutionary Acts beyond the defined framework of the spoken conversation and interaction. The distinct types of graphic patterns and visual representations contribute to the evaluation of spoken political and journalistic texts such as interviews, live conversations in the Media, as well as discussions in Parliament, focusing in the discourse component of spoken political and journalistic texts.

The graphic patterns and visual representations are based on the output of an interactive annotation tool for spoken journalistic texts presented in previous research [2]. Specifically, in the interactive annotation tool [2], incoming texts to be processed constitute transcribed data from journalistic texts. The annotation tool was designed to operate with most commercial transcription tools, some of which are available online. The development of the tool is based on data and observations provided by professional

journalists (European Communication Institute, Program M.A in Quality Journalism and Digital Technologies, Danube University at Krems, Athena- Research and Innovation Center in Information, Communication and Knowledge Technologies, Athens - Institution of Promotion of Journalism Ath.Vas. Botsi, Athens and the National and Technical University of Athens, Greece). Since processing speed and the option of re-usability in multiple languages of the written and spoken political and journalistic texts constitutes a basic target of the proposed approach, strategies typically employed in the construction of Spoken Dialog Systems, such as keyword processing in the form of topic detection, were adapted in the developed annotation tool. The functions of the designed and constructed interactive annotation tool [2] include providing 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. Furthermore, these functions facilitate the comparison between discourse structures of conversations and interviews with similar topics or the same participants/participant.

## 2 User Interaction and Relation Types

### 2.1 Interactive Registration of Relation Types

In the above-stated process of interactive annotation [1, 2], the “Identify Topic” command allows the content of answers, responses and reactions to be checked in respect to the question asked or issue addressed. Specifically, topics in respect to the question asked or issue addressed by the interviewer or moderator are defined at a local level with the activation of the “Identify Topic” command. Topics, treated as local variables, are registered and tracked. Assistance in choice of topic is provided to the user with the automatic signalization of nouns. Nouns are signalized by the Stanford POS Tagger in each turn taken by the speakers in the respective segment in the dialog structure [1]. The use of the registered and tracked keywords, treated as local variables, is crucial for the signalization of each topic and the relations between topics, since automatic Rhetorical Structure Theory (RST) analysis procedures [18, 26] usually involves larger (written) texts and may not produce the required results.

Relation types between topics are determined by the user by activating the “Identify Relation” command. We note that in the domain of journalistic texts, the relations between topics cannot be strictly semantic: automatic processes may result to errors. The user chooses among four available relations between the topic of the question or issue addressed with the topic of the respective response or reaction [2]: “Repetition”, “Association”, “Generalization” or “Topic Switch”.

The “Repetition” relation (“REP” tag) involves the repetition of the same word or synonym and corresponds to the generation of the shortest distance between defined topics, referred to as “Distance 1”. A characteristic example is “Britain”-“the UK” [1]. The “Association” relation (“ASOC” tag) is often defined by the user’s beliefs and world knowledge, for example, in the relation between “propaganda”-“social-media” [1]. The “Association” relation is represented as a longer line to the next word-node, corresponding to “Distance 2”. The generation of the longest distance between defined topics,

“Distance 3”, corresponds to the “Generalization” relation (“GEN” tag). The “Generalization” relation is also defined by the user’s world knowledge, however, in many cases, this relation can be evaluated with a lexicon or Wordnet, as in the example “police”-“security” [1]. The “Topic Switch” relation (“SWITCH” tag) is used when no evident semantic relations are identified between topics and the relation is perceived as a change (or switch) of topic, for example, in the relation between “security” and “entrepreneurship” [1]. “Topic Switch” generates a break in the sequence of topics [1, 2].

Observed differences between identified topic relations among some journalists that are non-native speakers of English [1] (especially in respect to “ASOC” and “SWITCH”) may in some cases be attributed to lack of world knowledge of the language community concerned [8, 15, 22]. As noted in previous research [1], these observations imply that the international public may often perceive and receive different and/or incomplete information in respect to evaluating conversation and interaction [3, 11, 16, 25].

## 2.2 Relation Types and Graphic Representation

The graphic pattern of the discourse structure concerning the interview or discussion is based on the representation of the selected local topics constituting the path of the user’s choices and interaction. In particular, path generation of the interaction is modeled and implemented based on user interactions registered in spoken dialog systems (in the domains of consumer complaints and mobile phone services call centers) [7, 14]. A visual representation from the user’s interaction is generated, tracking the corresponding selected keywords in the dialog flow. In the present application, the same model is applied for tracking topics and generating models in transcribed spoken journalistic texts [1, 2].

With the interactive generation of registered paths, similar to paths with generated sequences of recognized keywords [7, 14] a keyword (topic) may be repeated (“Repetition” relation) or related to a more general concept (or global variable) [10] (“Generalization” relation) or related to keywords (topics) concerning similar functions (“Association” relation). Similarly to the domain of spoken dialog systems, a keyword involving a new command or function is registered as a new topic (“Topic Switch” relation). Subsequently, the “path” of interaction is generated with the sequence of topics chosen by the user and the perceived relations between them. The generated “path” of interaction forms distinctive visual representations according to its content. Furthermore, topics and words generating diverse reactions and choices from different users may result to the generation of different forms of generated graphic representations for the same conversation or interaction.

## 3 Evaluation and Cognitive Bias

Generated graphic patterns contribute to a user-independent evaluation of spoken Human-Human conversation and interaction [1], similarly to user-independent evaluation of spoken dialog systems [23], where speed and correctness are of crucial importance [10]. In spoken dialog systems, varying degrees of user’s familiarity with

dialog systems or user-friendly interfaces in spoken interaction result to different perceptions of successful interactions. Thus, occasional errors may be “forgiven” by the user [6, 13]. Specifically, errors in spoken input or a longer duration of interaction due to complications in the dialog may not always correspond to negative evaluation. In a similar manner, varying degrees of familiarity and bias with topics discussed in spoken journalistic texts result to different perceptions of successful conversations or debates and any complications or mistakes can be “forgiven” by the user [1].

The content and form of the generated graphic representations can contribute in depicting the degree in which all topics are addressed as well as what topics are avoided. Topics introduced in the discussion or interview are avoided by speakers either by changing a topic or by persisting to address the same topic. The degree in which all topics are addressed, as well as what topics are avoided, are evident in the form of the generated graphic representation. For example, multiple breaks in the generated graphic representation correspond to multiple instances of topic switch and the (“New Topic” relation. Furthermore, the generated graphic representation may also depict how participants may be lead or even forced into addressing a topic – by association or generalization (the “Association” and the “Generalization” relations respectively) [1].

The content and form of the generated graphic representations (presented in the following section) may be considered as visual representations of Cognitive Bias [1], where the perceived relations-distances between word-topics perceived by the user are related to Lexical Bias [19]. Additionally, the graphic representations allow the determination of the participants in the conversation or interview who were successful in their spoken interaction and the participants who were less successful. This output targets to by-pass Confidence Bias [9] of users-participants and evaluators [1].

## 4 Form of Generated Graphic Representations

### 4.1 Present Approach

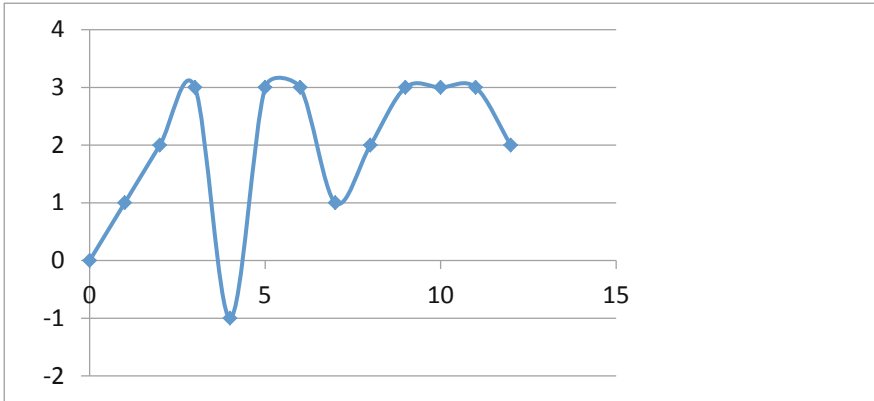
As described above, the generated graphic representation is based on the relations of the topics to each other, including distances from one word to another. In previous research [1, 2], Distances 1, 2 and 3 were depicted as vertical lines from top to bottom, in the case of the generation of a tree-like structure, or as horizontal lines from left to right, in the case of the generation of a graph. Topic switches were depicted as breaks in the continuous flow of the generated graphic representation, generating a new, disconnected point or node. This approach envisioned a possible further development with graphic forms similar to discourse trees [5, 12], however, it presented difficulties in matching points of the generated structure to the respective segments of the spoken text.

The present approach targets to allow the alignment of the generated graphic representation with the respective segments of the spoken text, facilitating a possible integration in transcription tools.

Similarly to the approaches presented in previous research [1, 2], the length of the lines between points corresponding to topics depends on the type of distance to the next word-node, with the shortest line corresponding to the relation of “Repetition”,

related to Distance 1 and the longest line corresponding to the relation of “Generalization”, Distance 3.

In the present application, Distances 1, 2 and 3 correspond to the respective values “1”, “2” and “3” ( $y = 1$ ,  $y = 2$  and  $y = 3$ ) depicted in the generated graphic representation. The “Topic Switch” relation is assigned value “-1” (Fig. 1).



**Fig. 1.** Distances and values between topics.

The starting point of the graphic representation of the spoken interaction depicted in Fig. 1 is point zero in the time frame (x), where  $(x, y) = (0, 0)$ . For the 1<sup>st</sup> second of spoken interaction there is an occurrence of two (2) keywords and one “Repetition” relation between them, represented as value “1” in the y axis (y), where (REP): 1, corresponding to point (1, 1).

From the 1<sup>st</sup> to the 2<sup>nd</sup> second ( $y = 2$ ) of spoken interaction, the 3<sup>rd</sup> keyword demonstrates an “Association” relation with the previous, 2<sup>nd</sup> keyword, represented as value “2” in the y axis (y), where (ASOC): 2, corresponding to point (2, 2).

Until the 3<sup>rd</sup> second of spoken interaction, there is one more 4<sup>th</sup> keyword and its relation with the previous, 3<sup>rd</sup> keyword is an “Generalization” relation, represented as value “3” in the y axis (y), where (GEN): 3, corresponding to point (3, 3).

In the 4<sup>th</sup> second of spoken interaction, the 5<sup>th</sup> keyword demonstrates a “New Topic” relation with the previous, 4<sup>th</sup> keyword, represented as value “-1” in the y axis (y), (NEW TOPIC): -1, corresponding to point (4, -1).

Two “Generalization” relations follow in the 5<sup>th</sup> second and 6<sup>th</sup> second of spoken interaction, where the relation between the 6<sup>th</sup> keyword and the previous, 5<sup>th</sup> keyword and the following 7<sup>th</sup> keyword is represented as value “3” in the y axis (y), where (GEN): 3, corresponding to points (5, 3) and (6, 3).

In the 7<sup>th</sup> second of spoken interaction, the 8<sup>th</sup> keyword there is a “Repetition” relation to the previous, 7<sup>th</sup> keyword, represented as value “1” in the y axis (y), where (REP): 1, corresponding to point (7, 1). In the 8<sup>th</sup> second of spoken interaction, the 9<sup>th</sup> keyword is related to the previous 8<sup>th</sup> keyword with an “Association” relation, represented as value “2” in the y axis (y), where (ASOC): 2, corresponding to point (8, 2).

A sequence of three “Generalization” relations follow in the 9<sup>th</sup> to 11<sup>th</sup> second of spoken interaction, where the relation between the 10<sup>th</sup> keyword and the previous, 9<sup>th</sup> keyword and the following 11<sup>th</sup> and 12<sup>th</sup> keywords is represented as value “3” in the y axis (y), where (GEN): 3, corresponding to points (9, 3), (10, 3) and (11, 3).

Finally, in the 12<sup>th</sup> second of spoken interaction, there is one more 13<sup>th</sup> keyword and its relation with the previous, 12<sup>th</sup> keyword is an “Association” relation between them (ASSOC): 2, corresponding to point (12, 2).

### 4.2 Graphic Representation and Relation Type

Dialog segments typically demonstrate a variety of topic relations, with a characteristic example shown in the above-described Fig. 1 and in Fig. 2. The typical variety of topic relations concerns all - or almost all- types of topic relations. Empirical data so far demonstrates a predominance of “Association” relations, a slightly lower occurrence of “New Topic” and “Generalization” relations and a low occurrence of “Repetition” relations. In the following examples (Figs. 2, 3, 4, 5 and 6) we present dialog segments of 12 s (12 s) with 13 word-topics and 12 relations between each word-topic, where x = time in secs and y = relation between two topics.

The example in Fig. 2 depicts two (2) “New Topic” relations (NEW TOPIC), corresponding to Distance  $y = -1$ , where there is a switch of topic (where  $y = -1$  and  $x = \{3, 12\}$ ). The example in Fig. 2 includes two (2) “Repetition” relations (REP) (where  $y = 1$ , και  $x = \{7, 10\}$ ), five (5) “Association” relations (ASOC) (where  $y = 2$  and  $x = \{2, 5, 8, 9\}$ ) and three (3) “Generalization” relations (GEN), where  $y = 3$  and  $x = \{1, 4, 6\}$ .

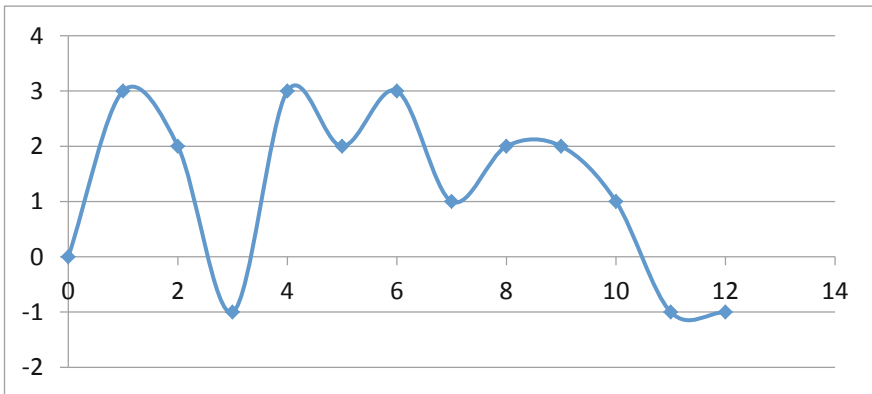


Fig. 2. Typical form of generated graphical representation.

### 4.3 Graphic Representation of “Repetition” Relations

In contrary to the examples presented in Figs. 1 and 2, a remarkable predominance of specific types of relations results to the generation of characteristic types of graphic representations. As previously described above, the overall shape of the generated

graphic representation is dependent on the mostly occurring relation types in the discourse structure of the interview or discussion.

A high frequency of “Repetition” relations is presented in Fig. 3, where seven (7) “Repetition” relations are registered with  $y = 1$  and  $x = \{1, 3, 5, 6, 7, 9, 11, 12\}$ . The same topic is repeated between the points in the above-presented  $x$  values.

The graphic representation in Fig. 3 demonstrates a development around the value  $y = 1$  level.

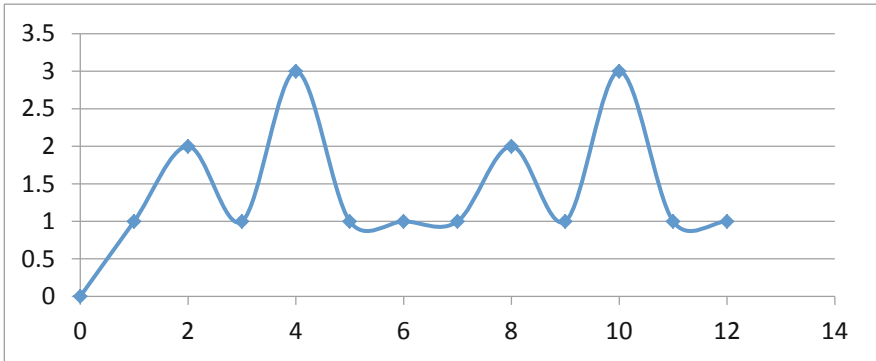


Fig. 3. Generated graphical representation with a “Repetition” relation.

#### 4.4 Graphic Representation of “Association” Relations

The generation of a graphic representation of multiple high peaks is illustrated in the example in Fig. 4, corresponding to transcripts of available online interviews. The characteristic plateau-like shape of the peaks in the generated graphic representation is affected by the relatively high percentage of “Association” relations on the value  $y = 2$  level. The present example (Fig. 4) depicts twelve (12) relations, several of which are “Association” (ASOC) relations, where  $y = 2$  and  $x = \{2, 5, 6, 8, 9, 10\}$ .

The graphic representation in Fig. 4 demonstrates a development around the value  $y = 2$  level.

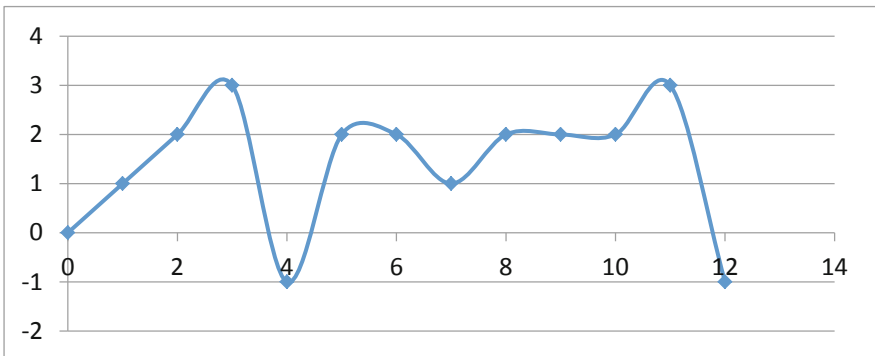


Fig. 4. Generated graphical representation with multiple “Association” relations.

### 4.5 Graphic Representation of “Topic Switch”

The generation of a graphic representation of many separate sharp peaks is illustrated in the following example in Fig. 5, corresponding to transcripts of available online interviews. In particular, the overall shape of the generated graphic representation is affected by the relatively high percentage of “Topic Switch” relations, creating a characteristic sequence of sharp peaks.

A high frequency of “New Topic” relations is presented in Fig. 4, where eight (8) “New Topic” relations are registered with  $y = -1$ , for  $x = \{1, 3, 5, 6, 7, 9, 11, 12\}$ . There is a change of topic between the points in the above-presented  $x$  values.

The graphic representation in Fig. 5 demonstrates multiple sharp drops in the value  $y = -1$  level.

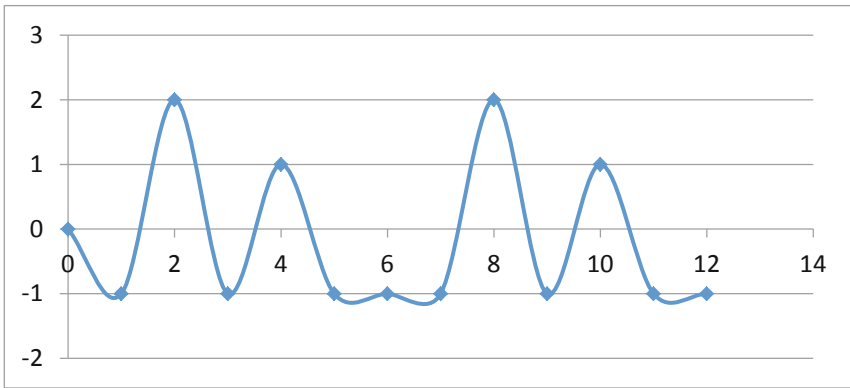


Fig. 5. Generated graphical representation with multiple “Topic Switch” relations.

### 4.6 Graphic Representation of “Generalization” Relations

The generation of a graphic representation of characteristically high peaks is illustrated in the example in Fig. 6, corresponding to transcripts of available online interviews. The characteristic plateau-like shape of the peaks in the generated graphic representation is affected by the relatively high percentage of “Generalization” relations on value  $y = 3$  level. The present example (Fig. 6) depicts six (6) “Generalization” (GEN) relations, where  $y = 3$  and  $x = \{3, 4, 5, 6, 9, 10, 11\}$  in which the “Generalization” (GEN) relation is repeated six (6) times.

The graphic representation in Fig. 6 demonstrates a development around the value  $y = 3$  level.



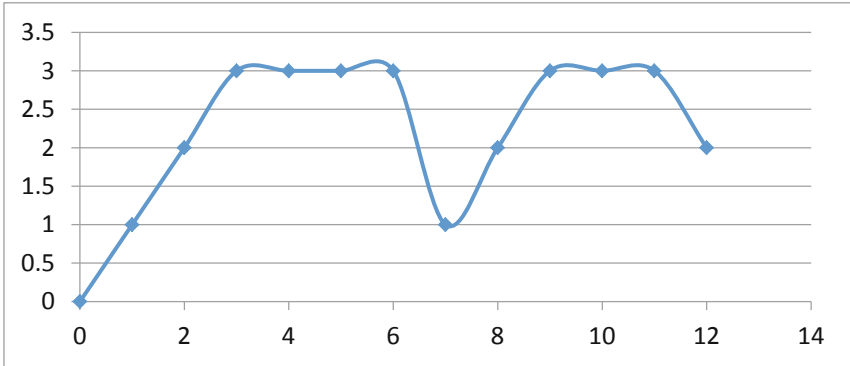


Fig. 6. Generated graphical representation with multiple “Generalization” relations.

## 5 Detecting Pointers to Speaker Intentions

The above-described graphic representations and values enable the evaluation of the behavior of speakers-participants, depicting Cognitive Bias and may also serve for bypassing Confidence Bias of the user-evaluator of the recorded and transcribed discussion or interview. Furthermore, the above-described graphic representations and values also allow the identification and detection of additional, “hidden” Illocutionary Acts not restricted to “Obtaining Information Asked” or “Providing Information Asked”, as defined by the framework of the interview or discussion.

Speech Acts performed by one or multiple speakers-participants usually involve complex Illocutionary Acts beyond the defined framework of the interaction. This feature differentiates Speech Acts in two-party or multiparty discussions or interviews from task-specific dialogs [20] and typical collaborative dialogs [21, 24]. In two-party or multiparty discussions or interviews, the Illocutionary Act [4, 17] performed by the Speakers may not be restricted to “Obtaining Information Asked” or “Providing Information Asked” in the spoken interaction concerned and may involve other or additional intentions regarding the presence of the speakers-participants and their role in the interview or discussion. In particular, the Illocutionary Acts not restricted to “Obtaining Information Asked” or “Providing Information Asked” may be related to one or more categories of Speech Acts defining less explicitly expressed Speaker intentions. Here, we present three frequently detected categories of pointers to “hidden” Speech Acts, namely “Presence”, “Express Policy” and “Make Impression”. We note that all three Speech Act pointers may be connected to each other and may even occur at the same time. The “Make Impression” Speech Act pointer is distinguished from the other two Speech Act pointer since it is identifiable on the Paralinguistic Level.

**The “Presence” Pointer.** “Presence” Speech Act pointer is identified by the Speaker’s reluctance to answer questions, avoidance of topics, or a polite or symbolic presence in the discussion or interview but not an active participation. Besides the Speaker’s silence (Silence/No Answer) as response to questions or statements, a “Presence” pointer is signaled by remaining in the same “safe” topic by repeating the same

subject (“Repetition”) or by introducing a “safer” and more general topic (“Generalization”) or a different topic (“Topic Switch”). “Presence” Speech Act pointers can be identified by a high frequency of one or more of the above-described relations, especially in combination with instances of no response (Silence/No Answer).

**The “Express Policy” Pointer.** With the “Express Policy” additional “hidden” Speech Act pointer, there is a direct or even blatant expression of opinion or policy. In this case, the Speaker may persist on discussing the same topic of interest by repeating the same subject (“Repetition”) or may try to direct the discussion in the topic(s) or interest by “Topic Switch”. Unlike the “Presence” pointer, the “Express Policy” pointer is characterized by a higher level of complexity, since it may contain features of the “Presence” Speech Act pointer and features of the “Make Impression” Speech Act pointer. However, in contrast to the case of the “Presence” pointer, the repeated topic(s) or the topics introduced are all - or almost all- semantically or associatively related. Although, as previously described, the “Express Policy” pointer may be related to the “Make Impression” Speech Act pointer, the “Express Policy” Speech Act pointer does not necessarily entail the creation of tension in the discussion or interview.

**The “Make Impression” Pointer.** With the “Make Impression” Speech Act pointer, the Speaker purposefully creates tension in the interview or discussion. The “Make Impression” pointer is characterized by any of the features of the “Presence” or “Express Policy” Speech Acts pointer. Additionally, the “Make Impression” pointer can also be distinguished from the previous Speech Act pointers in respect to features in the paralinguistic level of one (or all) of the Speakers, including rise of amplitude, prosodic emphasis and other prosodic features, gestures and facial expressions.

## 6 Conclusions and Further Research

The present application targets to assist evaluation and decision-making process in respect to discussions and interviews in the Media, providing a graphic representation of the discourse structure and aiming to by-pass Cognitive Bias of the user-evaluator. The predominate types of relations, if applicable, are easily identified by the y level value around which the graphic representation is developed.

The time frame generation of the linear structure allows the graphic representation to be presented in conjunction with the parallel depiction of speech signals and transcribed texts, a typical feature of most transcription tools.

Furthermore, the above-described graphic representations and values enable the evaluation of the behavior of speakers-participants, allowing the identification and detection of additional, “hidden” Illocutionary Acts not restricted to “Obtaining Information Asked” or “Providing Information Asked” framework defined by the interview or discussion. In the light of the above, the present application may also be adapted to additional domains, such as education-training and virtual negotiators, since it concerns the evaluation of a user’s familiarity, perception and world knowledge. The alignment of the generated graphic representation with the respective segments of the spoken text enable a possible integration of the present application in existing transcription tools.

## References

1. 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 (2018)
2. 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](https://doi.org/10.1007/978-3-319-21006-3_26)
3. Alexandris, C.: English, German and the international “semi-professional” translator: a morphological approach to implied connotative features. *J. Lang. Transl.* **11**(2), 7–46 (2010)
4. Austin, J.L.: *How to Do Things with Words* (eds. by, J.O. Urmson, M. Sbisà), 2nd edn. University Press, Oxford Paperbacks, Oxford (1962/1976)
5. 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)
6. Cohen, P., et al.: Quickset: multimodal interaction for distributed applications. In: *Proceedings of the 5th ACM International Multimedia Conference*, Seattle, Washington, pp. 31–40. Association for Computing Machinery (ACM) (1997)
7. 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)
8. Hatim, B.: *Communication Across Cultures: Translation Theory and Contrastive Text Linguistics*. University of Exeter Press, Exeter (1997)
9. Hilbert, M.: Toward a synthesis of cognitive biases: how noisy information processing can bias human decision making. *Psychol. Bull.* **138**(2), 211–237 (2012)
10. 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)
11. 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)
12. 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 (1999)
13. Nass, C., Brave, S.: *Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship*. The MIT Press, Cambridge (2005)
14. Nottas, M., Alexandris, C., Tsopanoglou, A., Bakamidis, S.: A hybrid approach to dialog input in the citizenship dialog system for consumer complaints. In: *Proceedings of HCI 2007*, Beijing, People’s Republic of China (2007)
15. Paltridge, B.: *Discourse Analysis: An Introduction*. Bloomsbury Publishing, London (2012)
16. Pan, Y.: Politeness in Chinese Face-to-Face Interaction. *Advances in Discourse Processes Series*, vol. 67. Ablex Publishing Corporation, Stamford (2000)
17. Searle, J.R.: *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, Cambridge, MA (1969)
18. Stede, M., Taboada, M., Das, D.: *Annotation Guidelines for Rhetorical Structure*. Manuscript. University of Potsdam and Simon Fraser University, March 2017 (2017)
19. Trofimova, I.: Observer bias: an interaction of temperament traits with biases in the semantic perception of lexical material. *PLoS ONE* **9**(1), e85677 (2014)

20. Tung, T., Gomez, R., Kawahara, T., Matsuyama, T.: Multi-party human-machine interaction using a smart multimodal digital signage. In: Kurosu, M. (ed.) *Human-Computer Interaction. Interaction Modalities and Techniques*, Lecture Notes in Computer Science, vol. 8007, pp. 408–415. Springer, Heidelberg (2013). [https://doi.org/10.1007/978-3-642-39330-3\\_43](https://doi.org/10.1007/978-3-642-39330-3_43)
21. Wang, H., Gailliot, A., Hyden, D., Lietzenmayer, R.: A knowledge elicitation study for collaborative dialogue strategies used to handle uncertainties in speech communication while using GIS. In: Kurosu, M. (ed.) *HCI 2013. LNCS*, vol. 8007, pp. 135–144. Springer, Heidelberg (2013). [https://doi.org/10.1007/978-3-642-39330-3\\_15](https://doi.org/10.1007/978-3-642-39330-3_15)
22. Wardhaugh, R.: *An Introduction to Sociolinguistics*, 2nd edn. Blackwell, Oxford (1992)
23. 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, 30 July–4 August 2017, pp. 665–677. Association for Computational Linguistics (ACL) (2017)
24. Yang, Z., Levow, G.A., Meng, H.: Predicting user satisfaction in spoken dialog system evaluation with collaborative filtering. *IEEE J. Sel. Top. Signal Process.* **6**(8), 971–981 (2012)
25. 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)
26. 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>