Processing Information Unspoken: New Insights from Crowd-Sourced Data for Sentiment Analysis and Spoken Interaction Applications

Christina Alexandris [1]

[1] *National and Kapodistrian University of Athens, Zografou Campus, Athens, Index, Greece*

Abstract

Crowd-sourced data offers new insights in the processing of information not uttered in spoken interaction. This subjective, perceived, context-related information, and its conversion into "visible" information in knowledge graphs for use in vectors / other forms of training data contributes to registering complex emotions in Sentiment Analysis, to monitoring fairness in spoken interaction and to data enrichment in HCI/HRI applications. Additionally, insights from crowd-sourced data allow a differentiation between circumstantial factors / evidence and socio-culturally-biased factors /evidence in data analysis and training data.

Keywords

Knowledge Graphs, Crowd-Sourced Data, Sentiment Analysis, Cognitive Bias, Plutchik Wheel of Emotions, Human-Computer Interaction, Spoken Dialog Systems

1. Fairness, Bias and Unspoken Information

Crowd-sourced data allows new insights in the analysis and processing of information not uttered in spoken interaction and the conversion of this information into "visible" and processable information in the form of knowledge graphs. The knowledge graph data, with subsequent use in vectors and other forms of training data [1] [2] [3] [4] are intended, at least in the present stage, as a dataset for training a neural network, with the possibility of conversion in Graph Neural Networks [5]. The conversion of knowledge graphs into training data contributes to the integration and processing of complex information and information not uttered in Natural Language Processing (NLP) tasks, thus, contributing to the creation of even more sophisticated systems. This possibility would not be considered if the above-stated characteristic research work were not accomplished.

The very nature and structure of knowledge graphs allows the representation of multiple facets of information – the multiple facets of the "Sense" of the words and/or transcribed video speech segments – although it is considered that there may exist some types of information/ some cases that may not have 100% coverage by a knowledge graph.

The detecting and processing of information not uttered but perceived-sensed by speakers-participants allows the integration of additional information content – meanings/senses- in training data. This allows the enrichment of data and a deeper understanding of speaker-participant psychology-mentality and sensitivities, contributing to a deeper understanding of the possible impact or consequences of a spoken journalistic/political text or interview or a video in Social Media (a). This also allows an additional approach to registering of cause-result relations on a discourse basis, including the monitoring of Fairness, namely that all voices-aspects-opinions are heard clearly –that all participants are given a fair chance in the interview or discussion and are

AAAI 2023 Spring Symposia, Socially Responsible AI for Well-being, March 27–29, 2023, USA
EMAIL: calexandris@gs.uoa.gr (A. 1)

not purposefully or unconsciously repressed, oppressed, offended or even bullied (b). The way sensitive topics and speaker-participant sensitivity are purposefully or unconsciously treated and managed contributes to registering and monitoring fairness in spoken interaction, avoiding Confidence Bias [6]. In particular, a crucial element in achieving "visibility" and, subsequently, "processability" of information not uttered is causality, namely the registration and processing of reactions triggered by that very information not uttered - the multiple facets of the "Sense" of the words in transcribed video and speech segments and in Social Media.

These reactions include subtle negative reactions in the Plutchik Wheel of Emotions, namely "Apprehension", "Annoyance", "Disapproval", "Contempt", "Aggressiveness" [7] - emotions usually too subtle to be easily extracted by sensor and/or speech signal data [8] [9] [10]. Additionally, the detecting and processing of information not uttered (often emotionally "sensitive" information) contributes in Sentiment Analysis (and Opinion Mining) applications where spoken data and/or videos are processed. However, crowd-sourced input indicates that information not uttered, along with subtle emotions – occurring in the outer circles - of the Plutchik Wheel of Emotions, may be (1) differently (or falsely) perceived – especially by non-native speakers of a natural language", (2) may be highly dependent on random and/or circumstantial or individual-specific factors and (3) may concern specific domains and related discourse. For Sentiment Analysis (and Opinion Mining) applications, (1), (2) and (3) are equally important.

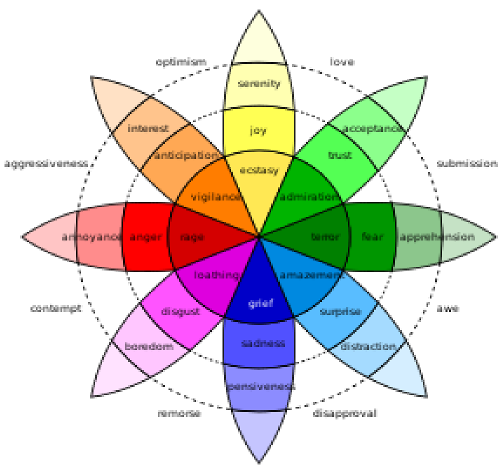

Figure 1: The Plutchik Wheel of Emotions

Considering the targets (a) and (b) contributing to "Socially Responsible AI", the purpose of the present approach is to account for the new (complex) factors/ insights gained from crowd-sourced data and their integration into knowledge graphs with subsequent use in training data - neural networks. The main focus is on the data preparation stage for subsequent extensive implementation and quantitative evaluation.

2. Processing Unspoken Information and Knowledge Graphs- the "Context" Relation

The knowledge graphs, generated by an interactive application presented in related/previous research [11] [12] [13], involve the depiction of two main categories of information not uttered in spoken interaction:

(I) Additional perceived information content and dimensions of –notably- very common words – information not registered in language resources, it may concern context-specific socio-cultural associations and Cognitive Bias, in particular, Lexical Bias [14]. (II) Perceived paralinguistic elements influencing the information content of spoken utterances. Both types of perceived information are language- and socio-culturally specific and are purposefully or subconsciously conveyed or perceived-understood by speakers-participants in the same language community.

In the knowledge graphs, this additional information of the above-described categories (I) and (II) is linked as an additional node to the spoken word with the proposed "Context" relation. The knowledge graphs can, subsequently, be converted into vectors and other forms of training data which is targeted to contain (a) "visible" and processable information not uttered in spoken interaction and (b) multiple versions and varieties of training data with perceived information generated by the implemented interactive application [11] [12] [13].

In our previous research [12] [15] [16], 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.

Although the concept of the generated graphic representations originates from the Discourse Tree prototype [17], the characteristics of spontaneous turn-taking [18] and short spoken speech segments did not facilitate the implementation of typical strategies based on Rhetorical Structure Theory (RST) [19] [20] [21].

In particular, strategies typically employed in the construction of most Spoken Dialog Systems were adapted in an interactive annotation tool designed to operate with most commercial transcription tools [12] [15] [13]. These strategies include keyword processing in the form of topic detection from which approaches involving neural networks are developed [22] [23]. The output provides the User-Journalist with (i) the tracked indications of the topics handled in the interview or discussion and (ii) the graphic pattern of the discourse structure of the interview or discussion. The output (i) and (ii) 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) [13] as well as the degree of tension in their interaction ("TENSION" Module). These features are identified by a set of criteria based on the Gricean Cooperative Principle [24] [25] (including paralinguistic elements). The implemented "RELEVANCE" Module [13] is intended for the evaluation of short speech segments and 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. This concerns topics avoided, introduced or repeatedly referred to by each Speaker-Participant (Repetitions, Associations, Generalizations and Topic Switches). The assigned respective values of each relation ("Relevance (X)" benchmark, [16]) were converted into generated visual representations and were registered as tuples or as triple tuples and, subsequently, converted into knowledge graphs.

In the context of spoken interaction, Cognitive Bias may concern "Association" (or other) relations and argumentation related to inherent yet subtle socio-culturally determined linguistic features in (notably) commonly occurring words presented in previous research (examples from the international community: (the) "people", (our) "sea"). These word types are detectable from the registered reactions [26] they trigger in the processed dialog segment with two (or multiple) speakers-participants. Since these words are very common and do not contain descriptive features, the subtlety of their content is often unconsciously used or is perceived (mostly) by native speakers and may contribute to the degree of formality or intensity of conveyed information in a spoken utterance. Here, these words concerning Cognitive Bias – Lexical Bias are referred to as "Gravity" words [26]. In other cases, these word types, although common words, may contribute to a descriptive or emotional tone in an utterance and they may play a remarkable role in interactions involving persuasion and negotiations. Specifically, it is considered that, according to Rockledge et al, 2018 [27], "the more extremely positive the word, the greater the probability individuals were to associate that word with persuasion". Here, these words concerning Cognitive Bias – Lexical Bias are referred to as "Evocative" words [26]. The subtle impact of words is one of the tools typically used in persuasion and negotiations [28] [29].

Generated graphical representations of perceived word-topic relations and registered "Gravity" and "Evocative" words (concerning Cognitive Bias – Lexical Bias) can be converted into sequences for their subsequent conversion into knowledge graphs or other forms of data for neural networks and Machine Learning applications [1] [2] [3] [4]. As described in previous research [15], registered "Gravity" and "Evocative" words are appended as marked values with "&" in the respective tuples or triple tuples. In the sequences with the respective tuples or triple tuples, the "&" indication is converted into a "CONTEXT" relation. In the knowledge graphs, this additional information is linked as an additional node to the spoken word with the proposed "Context" relation. The term "Context" is chosen to signalize the perceived context of additional information in the form of co-occurring linguistic and/or paralinguistic features, influencing the information content of the spoken utterance and its impact in the spoken interaction and dialogue structure.

In the case of paralinguistic elements, the "Context" relation links an additional expression – a word-entity, to the word uttered, for example, a modifier [30], completing its perceived content. This practice is typical of professional translators and interpreters when correctness and precision is targeted [31], as research and reports demonstrate. The "CONTEXT" relation connects the chosen word-topic from the speech segment with a word-expression emphasizing / complementing the spoken content such as "important" or a respective word summarizing the message. We note that the "CONTEXT" relation may link both a "Gravity"/ "Evocative" word and a paralinguistic element to the word-topic of a spoken utterance (Fig. 2).

For paralinguistic features depicting contradictory information to the information content of the spoken utterance, the "CONTEXT" relation connects the chosen word-topic from the speech segment with a word-expression contradicting the spoken content with the expression "not really" as a special indication.

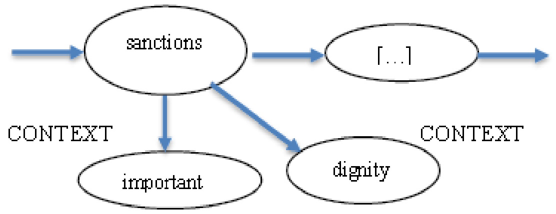

Figure 2: Fragment of knowledge graph for perceived meaning of eyebrow-raise ("important") co-occurring with topic "sanctions" and perceived "Gravity" word ("dignity") in spoken utterance.

3. New Insights for Knowledge Graphs and the Information (Atmo) "Sphere" of Spoken Words

The generated knowledge graphs from the interactively created visual representations for the same conversation and interaction may be compared to each other and be integrated in a database currently under development. Chosen relations between topics may describe Lexical Bias [14] and may differ according to political, socio-cultural and linguistic characteristics of the user-evaluator. This especially applies for international speakers/users [32] [33] [34] [35] [36], due to lack of world knowledge of the language community involved [37] [38]. In this case, it is considered that the registration of spoken interaction is dependent on user's perception and linguistic parameters and socio-cultural norms. This allows for a finite set of data to be pre-defined for evaluation and comparison and/or it used as seed data for the enrichment of existing data sets. However, with the extended integration of crowd-sourced input, the use of seed data for the enrichment of existing data sets does not apply in all cases. In particular, crowd-sourced input indicates that:

Unspoken Information may be differently (or falsely) perceived – especially by non-native speakers of a natural language - and especially when subtle emotions in the Plutchik Wheel of Emotions, are concerned (1).

Another important factor is that the perception of information not uttered may be highly dependent on random and/or circumstantial or individual-specific factors (2) or the perception of unspoken information may concern only specific domains and related discourse (3).

User-specific and crowd-sourced data may be problematic due to a number of factors concerning users' perception but also users' experience and time and effort invested in providing quality data – especially when very subtle linguistic and paralinguistic features are concerned. Therefore, it is necessary for the above-described problematic aspects of user-specific and crowd-sourced data to be minimized and/or controlled.

These observations from crowd-sourced data call for a differentiation between perceived unspoken information compatible to language-specific and socio-cultural norms and perceived unspoken information that is either strictly circumstantial or strictly domain/context dependent. Context-specific unspoken additional dimensions of individual spoken words may be described as an information (atmo) "sphere" surrounding the word, with the semantic content of the word in its nucleus, its context-specific and language-specific dimensions in the inner layer of the sphere (A) and its context-specific and non-language-specific dimensions in the outer layer of the "sphere" (B).

In other words, the actual semantic content of the word as defined in dictionaries and lexica (and hence, retrievable and processable) constitutes the center-nucleus of the "sphere" and is context-independent. The perceived unspoken context-specific dimensions of the word that are dependent on the above-described linguistic parameters and socio-cultural norms (such as "Gravity" and "Evocative" words and distinctive meanings of paralinguistic features) constitute the inner layer of the "sphere" (A). As previously mentioned, this information can constitute a finite set of pre-defined (seed) data for the enrichment of existing data sets, according to the type(s) of natural language(s) involved. This information may be not perceived or incorrectly perceived by non-native speakers-participants or by inexperienced speakers-participants due to age or training/background (i.e. crowd-sourced data from teenagers, users not familiar with i.e. sophisticated political speech) (i).

The perceived unspoken context-specific and non-language-specific dimensions of the word constitute the outer layer of the "sphere" (B). These non-language-specific dimensions account for information perceived by an individual as an isolated case or due to random and/or circumstantial factors of the current context (i).

The differentiation between context-specific dimensions of a spoken word that are language-specific and non-language-specific allows a differentiation between circumstantial factors/evidence and socio-culturally-biased factors/evidence in data analysis and training data.

The outer layer of the "sphere" also accounts for unspoken and non-language-specific dimensions of a word that are, however, domain-specific and/or related to a domain-specific discourse. For example, the word "follower" may be linked to different associations and subsequent dimensions of meanings and responses within a social media domain or within a geopolitical – war domain (ii). Furthermore, a word not expressing sentiment/emotion may be related to domain-specific positive or negative statements as observed in Sentiment Analysis and Opinion Mining applications [22]. A typical case are words that do not express sentiment but are connected to positive or negative statements as registered in Sentiment Analysis and Opinion Mining. For example, in restaurant reviews, the word "waiter" often occurs in negative statements [22].

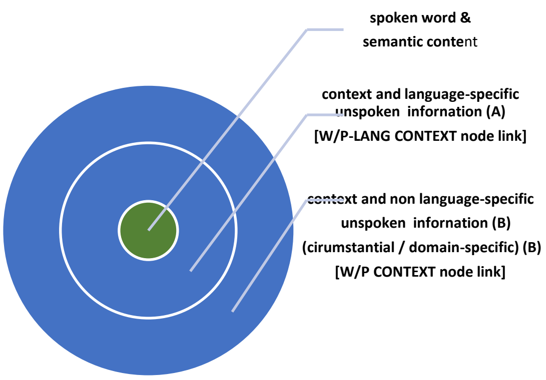

Figure 3: The (spoken) word (Atmo) "Sphere" of linguistic and paralinguistic information not uttered and respective types of "CONTEXT" relations in knowledge graphs.

The differentiation between different types of perceived unspoken information can be linked to the "CONTEXT" relation described in previous research, where the different types of perceived information not uttered can be differentiated with distinct types of "CONTEXT" relations in the knowledge graphs.

The context-specific and language-specific (A) "CONTEXT" relations employed in knowledge graph-based data are, henceforth, referred to as "W-LANG" CONTEXT relations for linguistic information not uttered, such as the additional content of "Gravity" and "Evocative" words. Additionally, the context-specific and language-specific "CONTEXT" relations are, henceforth referred to as "P-LANG" CONTEXT relations for paralinguistic information not uttered such as the above-described perceived meaning of a facial expression ("eyebrow-raise") related to language-specific and socio-cultural norms.

The non-language-specific /domain-specific (B) "CONTEXT" relations are, henceforth referred to as "W" CONTEXT relations for linguistic information not uttered and "P" CONTEXT relations for non-language-specific/ /domain-specific paralinguistic information.

4. Language-/Socio-culturally Specific Unspoken Information

In the proposed knowledge graphs (Fig. 4), language-specific dimensions in the inner layer of the sphere (A) include "Gravity" or "Evocative" words perceived by native speakers of a natural language that can be expressed with the W-LANG CONTEXT relation. The standard types of messages and information (and their variants) conveyed by paralinguistic features perceived by native speakers of a natural language can be expressed with the P-LANG CONTEXT relation (Fig. 5).

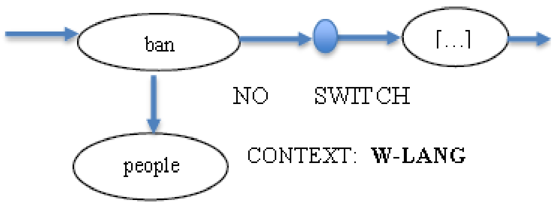

Figure 4: Fragment of knowledge graph for perceived "Evocative" word ("people"), co-occurring with topic "ban" resulting to a "No" answer and topic switch (SWITCH) in utterance segment with detected tension between speakers: context-specific and language-specific "CONTEXT: W-LANG" relation for linguistic information.

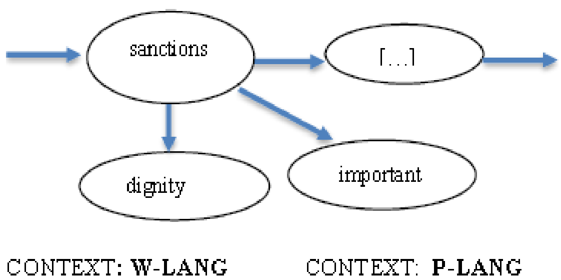

Figure 5: Fragment of knowledge graph for perceived meaning of eyebrow-raise

("important") co-occurring with topic "sanctions" and perceived "Gravity" word ("dignity") in utterance: context-specific and language-specific "CONTEXT: W-LANG" relation for linguistic information and context-specific and language-specific "CONTEXT: P-LANG" relation for paralinguistic information.

In regard to the language and culture-specific (standard) types of messages and information (and their variants) conveyed by paralinguistic features, examples of (interactively) annotated paralinguistic features depicting information complementing the information content of the spoken utterance are the following [26], for example: "[+ facial-expr: eyebrow-raise]" and "[+ gesture: low-hand-raise]") or constituting "stand-alone" information [26]. In the latter case, information was interactively annotated with the insertion of a separate message or response [Message/Response]. For example, the raising of eyebrows with the interpretation "I am surprised" [and / but this surprises me] [26] was indicated as [I am surprised] (a), either as a pointer to information content or as or as a substitute of spoken information, a "stand-alone" paralinguistic feature [Message /Response: I am surprised] [26]. Alternative interpretations of the paralinguistic feature are "I am listening very carefully" (b), "What I am saying is important"(c), "I have no intention of doing otherwise" (d) [26], indicated with the respective annotations according to the parameters of the language(s) and the speaker(s) concerned.

This type of (annotated) data for paralinguistic features constituting unspoken information may contribute to the management of problematic input in typical Data Mining and Sentiment Analysis-Opinion Mining applications, especially if the semantic content of a spoken utterance is complemented or contradicted by a gesture, facial expression, movement – or even by tone of voice. Typical Data Mining and Sentiment Analysis-Opinion Mining applications mostly rely on word groups, word sequences and/or sentiment lexica [39], including recent approaches with the use of neural networks [40] [41] [42].

As previously mentioned, with the present approach, this type of language-specific data – linguistic features and paralinguistic features- can be used as seed data for Sentiment Analysis and related applications. It can also be used as a base-line for comparison and evaluation of multiple user-input, especially if the quality of the crowd-sourced data is not guaranteed. The language-specific (seed) data can also be integrated in HCI applications intended for native or near-native speakers of a particular natural language or for a defined pair or set of languages.

5. Unspoken Non-Language Specific and Domain-Specific Information

In the case of non-language-specific information that is, however, domain-specific (B), the data can be integrated in domain-specific applications. For example, in Sentiment Analysis applications for restaurant reviews, the emotionally neutral words "bill" or "waiter" are connected with the dimension-meaning of a negative statement [22] with the "CONTEXT: W" relation (Fig.6). In other words, a positive or negative dimension may be automatically related to a word, depending on context – a feature of crucial importance in Sentiment Analysis.

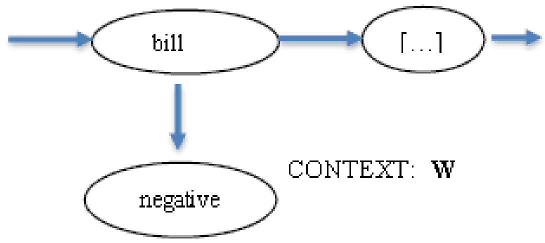

Figure 6: Fragment of knowledge graph for perceived word: context-specific and non-language-specific "CONTEXT: W" relation for linguistic information in the domain of "Restaurant Reviews" for Sentiment Analysis applications. The word "bill" is marked with a negative attitude.

The non-language-specific but strictly context-specific dimension of a word can also be domain-specific. For example, a particular word may imply a specific role or action. It may be noted that this allows possible implementations within a "frame-slot" framework in domain-specific (HCI and HRI) for processing spoken utterances. In this case, the mere utterance of a single word may imply a domain-specific type of information consisting a complete phrase or sentence – or one or more possible domain-specific alternative types of implied information (Fig. 7).

A characteristic example of non-language-specific features comprising additional dimensions of information content of words is the case of specific words receiving prosodic emphasis within the discourse and/or domain of the spoken interaction. Prosodic emphasis may stress and/or clarify the semantic content of the spoken utterance in a broad range of interaction types. These interaction types range from task-specific dialogue and question-answer interactions to interviews, political discussions and spoken interaction concerning negation and persuasion and/or expression of opinion.

Prosodic emphasis, change of tone of voice and speaker/individual-specific paralinguistic features can be inserted as additional information with the "Context" relation, as in the case of language-specific paralinguistic features presented in previous research. The context-specific and language-specific "W-LANG" CONTEXT and "P-LANG" CONTEXT relations for linguistic and paralinguistic information not uttered can be integrated with non-language-specific / domain-specific "W" CONTEXT and "P" CONTEXT relations for linguistic and paralinguistic information within a knowledge graph (Fig. 8). All types of linguistic and/or paralinguistic CONTEXT relations may co-occur within the same speech segment, although this is not considered common.

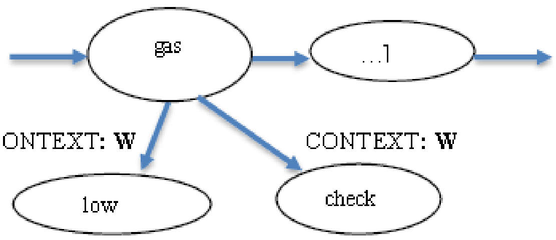

Figure 7: Fragment of knowledge graph for a singular spoken word "gas" and context-specific and non-language-specific "CONTEXT: W" relation for domain-specific information in HCI applications. The word "gas" is marked with implied possible information "(the [gas] is) low" and "check [gas]".

We note that the W-LANG and P-LANG "CONTEXT" relation or the W and P "CONTEXT" relations may be selected and be processed separately, according to application type. For example, language-specific data – linguistic features and paralinguistic features- can be used as seed data in a database –resource for language-specific applications. Non-language specific/domain-specific data – linguistic features and paralinguistic features- can be used as seed data in a separate database –resource for domain-specific applications. Both databases – resources can be merged according to application type. We note that recent approaches in estimating node importance in knowledge graphs may enable the automatic execution of such processes [43], however, further research is required.

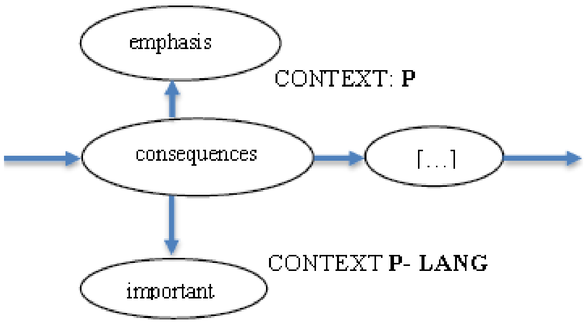

Figure 8: Fragment of knowledge graph for perceived word: language-specific (P-LANG) (eyebrow raise = "important") and non-language-specific (P) (prosodic emphasis) "CONTEXT" relations for paralinguistic information.

6. Conclusions and Further Research

Crowd-sourced data resulted to new insights in the analysis and processing of information not uttered in spoken interaction [44] and its integration in knowledge graphs, with its subsequent use in vectors and other forms of training data as dataset for training a neural network for Natural Language Processing (NLP) tasks. Insights from crowd-sourced data enabled a differentiation between perceived linguistic and paralinguistic information not uttered compatible to language-specific and socio-cultural norms and unspoken perceived information that is either strictly circumstantial or strictly domain/context dependent. This enables a differentiation between circumstantial factors/evidence (individual/context-specific or domain specific - for Sentiment Analysis/HCI) and socio-culturally-biased factors/evidence in data analysis and training data and its integration in knowledge graphs (1). In the latter case, language/socio-culturally-specific factors are more likely to account for speaker-participant psychology-mentality and sensitivities and for cases of intended or unintended offense or bullying, differentiating them from any random occurrences /individual-specific peculiarities (especially for paralinguistic features), thus, contributing to "Socially Responsible AI".

As proposed, context-specific additional dimensions of individual spoken words may be described as a context-specific information (atmo) "sphere" surrounding the spoken word. The concrete meaning – actual semantic content of the word (retrievable and processable in Natural Language Processing-NLP) is surrounded by two context-specific layers, with its context-specific and language-specific dimensions in the inner layer of the sphere (A) and its context-specific and non-language-specific dimensions in the outer layer of the "sphere" (B). The outer layers of the word (atmo) "sphere" demonstrate similarities to the outer circles of the Plutchik Wheel of Emotions containing complex emotions, recognizable within a (socio-culturally determined) context, such as "contempt" and "disapproval". In contrast, concretely identifiable emotions – including intense and universally recognizable emotions, such as "rage" and "grief" - are located in the inner circles of the Plutchik Wheel of Emotions and are typically easily detected and processed by current practices in Sentiment Analysis and Opinion Mining. In other words, the proposed information (atmo) "sphere" surrounding the spoken word mirrors the overall shape and very general – basic- features in the Plutchik Wheel of Emotions (2).

The distinct types of integration of the "Context" factor and related information in knowledge graphs –as provided by crowd-sourced data – outline the distinct types of implementation for the enrichment of models and refining NLP tasks — especially when videos and multimodal data are processed (3). In addition to their integration in knowledge graphs, the pre-defined words can also be used as an enhanced "Bag-of-Words" approach (Seed Data) in strategies and applications such as spoken Dialog Systems. In the case of Dialog Systems and related HCI/HRI applications, with the proposed processing strategy, the mere utterance of a single word may imply a complete phrase / sentence with domain-specific (alternative types of) information (4).

Since the present approach focuses on the data preparation stage, targeting to its contribution to "Socially Responsible AI", further research is geared towards the extensive implementation, evaluation (with quantitative evaluation measurements) and improvement of the training data created by the knowledge graphs, especially for a wider range of languages and speakers.

7. References

[1] S. Mittal, A. Joshi, T. Finin, Thinking, Fast and Slow: Combining Vector Spaces and Knowledge Graphs (2017) URL: arXiv:1708.03310v2 [cs.AI]

[2] M. Mountantonakis, Y. Tzitzikas, Knowledge Graph Embeddings over Hundreds of Linked Datasets, in: Garoufallou E., Fallucchi F., William De Luca E. (Eds.), Metadata and Semantic Research MTSR 2019, volume 1057 of Communications in Computer and Information Science, Springer, Cham, 2019, pp. 150-162. doi: 10.1007/978-3-030-36599-8_13

[3] H. N., Tran, A. Takashu, Analyzing Knowledge Graph Embedding Methods from a Multi-Embedding Interaction Perspective, in: Proceedings of the 1st International Workshop on Data Science for Industry 4.0 (DSI4) at EDBT/ICDT 2019 Joint Conference, 2019. URL: https://arxiv.org/abs/1903.11406

[4] M. Wang, L. Qiu , A Survey on Knowledge Graph Embeddings for Link Prediction, Symmetry, 13, 485 (2021). doi:10.3390/sym13030485

[5] Z. Ye, Y. J. Kumar, G. O. Sing, F. Song, J. Wang, A Comprehensive Survey of Graph Neural Networks for Knowledge Graphs, IEEE Access, 10 (2022). 75729-75741. doi: 10.1109/ACCESS.2022.3191784.

[6] M. Hilbert, Toward a Synthesis of Cognitive Biases: How Noisy Information Processing Can Bias Human Decision Making, Psychological Bulletin 138(2) March 2012 (2012) 211-237.

[7] R. Plutchik, A psychoevolutionary theory of emotions, Social Science Information 21 (1982) 529-553. doi: 10.1177/053901882021004003

[8] Z. He, T. Jin, A. Basu, J. Soraghan, G. Di Caterina, L. Petropoulakis, Human emotion recognition in video using subtraction pre-processing, in: Proceedings of the 2019 11th International Conference on Machine Learning and Computing, Zhuhai China, 2019, pp. 374–379.

[9] S. Poria, E. Cambria, D. Hazarika, N. Mazumder, A. Zadeh, L-P. Morency, 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, Association for Computational Linguistics – ACL, 2017, pp. 873–883. doi:10.18653/v1/P17-1081

[10] A. Yakaew, M. Dailey, T. Racharak, Multimodal Sentiment Analysis on Video

Streams using Lightweight Deep Neural Networks, in: Proceedings of the 10th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2021), 2021, pp. 442-451. doi: 10.5220/0010304404420451

[11] C. Alexandris, J. Du, V. Floros, Visualizing and Processing Information Not Uttered in Spoken Political and Journalistic Data: From Graphical Representations to Knowledge Graphs in an Interactive Application, in: M. Kurosu M. (Ed.), Human-Computer Interaction, Design and User Experience Case Studies, volume 13303 of Lecture Notes in Computer Science, Springer, Cham, 2022, pp. 211–226. doi:10.1007/978-3-031-05409-9_16

[12] C. Alexandris, V. Floros, D. Mourouzidis, Graphic Representations of Spoken Interactions from Journalistic Data: Persuasion and Negotiations, in: M. Kurosu (Ed.), Human-Computer Interaction, Design and User Experience Case Studies, volume 12764 of Lecture Notes in Computer Science LNCS, Springer, Cham, 2021, pp. 3-1. doi: 10.1007/978-3-030-78468-3_1

[13] D. Mourouzidis, V. Floros, C. Alexandris, Generating Graphic Representations of Spoken Interactions from Journalistic Data, in: M. Kurosu, (Ed.), volume 11566 of Lecture Notes in Computer Science LNCS, Springer, Basel, 2019, pp. 559–570.

[14] I. Trofimova, Observer Bias: An Interaction of Temperament Traits with Biases in the Semantic Perception of Lexical Material. PLoSONE 9(1): e85677 (2014).

[15] C. Alexandris, D. Mourouzidis, V. Floros, Generating Graphic Representations of Spoken Interactions Revisited: The Tension Factor and Information Not Uttered in Journalistic Data, in: M. Kurosu (Ed.) Human-Computer Interaction. Design and User Experience, volume 12181 of Lecture Notes in Computer Science, LNCS, Springer Nature, Switzerland, 2020, pp. 523–537. doi: 10.1007/978-3-030-49059-1_39

[16] C. Alexandris, Measuring Cognitive Bias in Spoken Interaction and Conversation: Generating Visual Representations, in: Beyond Machine Intelligence: Understanding Cognitive Bias and Humanity for Well-Being AI, Proceedings from the AAAI Spring Symposium, Stanford University, Technical Report, SS-18-03, Palo Alto, CA: AAAI Press, 2018, pp. 204-206.

[17] D. Marcu, Discourse trees are good indicators of importance in text, in: I. Mani, M. Maybury (Eds.), Advances in Automatic Text Summarization, The MIT Press, Cambridge, MA, 1999, pp. 123-136.

[18] M. Wilson, T.P. Wilson, An oscillator model of the timing of turn taking, Psychonomic Bulletin and Review 12 (6) (2005) 957-968.

[19] L. Carlson, D. Marcu, M. E. Okurowski, 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, 2001. URL:https://aclanthology.org/W01-1605.pdf

[20] M. Stede, M. Taboada, D. Das, Annotation Guidelines for Rhetorical Structure. Manuscript. University of Potsdam and Simon Fraser University, March 2017. URL: https://www.sfu.ca/~mtaboada/docs/research/RST_Annotation_Guidelines.pdf

[21] A. Zeldes, rstWeb - A Browser-based Annotation Interface for Rhetorical Structure Theory and Discourse Relations, in: Proceedings of NAACL-HLT 2016 System Demonstrations. San Diego, CA 2016, pp. 1-5. http://aclweb.org/anthology/N/N16/N16-3001.pdf

[22] D. Jurafsky, J. H. Martin, Speech and Language Processing, an Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition, 3rd. ed. Draft: https://web.stanford.edu/~jurafsky/slp3/ed3book_jan122022.pdf

[23] J.D. Williams, K. Asadi, G. Zweig, 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, Association for Computational Linguistics (ACL), 2017, pp. 665–677. URL: https://aclanthology.org/P17-1062/

[24] H. P. Grice, Studies in the Way of Words. Harvard University Press, Cambridge, MA 1989.

[25] H.P. Grice, Logic and conversation, in: P. Cole, J. Morgan, (Eds.), Syntax and Semantics, volume 3, Academic Press, New York (1975).

[26] C. Alexandris, Issues in Multilingual Information Processing of Spoken Political and Journalistic Texts in the Media and Broadcast News, Cambridge Scholars, Newcastle upon Tyne, UK, 2020.
[27] M.D, Rocklage, D.D. Rucker, L.F. Nordgren, Psychological Science 29(5) (2018) 749–760. doi: 10.1177/0956797617744797
[28] N. J., Evans, D. Park, Rethinking the Persuasion Knowledge Model: Schematic Antecedents and Associative Outcomes of Persuasion Knowledge Activation for Covert Advertising, Journal of Current Issues & Research in Advertising, 36:2 (2015) 157-176. doi: 10.1080/10641734.2015.1023873
[29] K. Skonk, 5 Types of Negotiation Skills, Program on Negotiation Daily Blog, Harvard Law School, May the 14th 2020. URL: https://www.pon.harvard.edu/daily/negotiation-skills-daily/types-of-negotiation-skills/ last accessed 2022/12/11.
[30] C. Alexandris, English, German and the International "Semi-professional" Translator: A Morphological Approach to Implied Connotative Features, Journal of Language and Translation, Sejong University, Korea, September 2010, 11 (2) (2010) 7- 46.
[31] W. Koller, Der Begriff der Äquivalenz in der Übersetzungswissenschaft, in: C. Fabricius-Hansen, J. Ostbo (Eds.), Übertragung, Annährung, Angleichung, Sieben Beiträge zu Theorie und Praxis des Übersetzens, Peter Lang, Frankfurt am Main, 2000, pp. 11-29.
[32] J. Du, C. Alexandris, D. Mourouzidis, V. Floros, A. Iliakis, Controlling Interaction in Multilingual Conversation Revisited: A Perspective for Services and Interviews in Mandarin Chinese, in: M. Kurosu (Ed.), volume 10271 of Lecture Notes in Computer Science LNCS, Springer-Verlag, Heidelberg, Germany, 2017, pp. 573–583.
[33] J. Ma, A comparative analysis of the ambiguity resolution of two English-Chinese MT approaches: RBMT and SMT, Dalian University of Technology Journal, 31(3) (2010) 114-119.
[34] B. Paltridge, Discourse Analysis: An Introduction, Bloomsbury Publishing, London, 2012.
[35] Y. Pan, Politeness in Chinese Face-to-Face Interaction, in: volume 67 of Advances in Discourse Processes series, Elsevier Science, Amsterdam, 2000.
[36] Z. W. Yu, Z. Y. Yu, H. Aoyama, M. Ozeki, Y. Nakamura, Capture, Recognition, and Visualization of Human Semantic Interactions in Meetings, in: Proceedings of PerCom, Mannheim, Germany, 2010, pp. 107-115.
[37] B. Hatim, Communication Across Cultures: Translation Theory and Contrastive Text Linguistics, University of Exeter Press, Exeter, UK, 1997.
[38] R. Wardhaugh, An Introduction to Sociolinguistics, 2nd. ed., Blackwell, Oxford, UK, 1992.
[39] B. Liu, Sentiment Analysis and Opinion Mining, Morgan & Claypool, San Rafael, CA, 2012.
[40] C. M. Arockiaraj, Applications of Neural Networks In Data Mining, International Journal Of Engineering And Science, volume 3, Issue 1 (May 2013), (2013) 8-11.
[41] M. A. Hedderich, D. Klakow, 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, Association for Computational Linguistics-ACL, 2018, pp. 12–18. https://aclanthology.org/W18-3402/
[42] K. Shah, S. Kopru, J-D. Ruvini, Neural Network based Extreme Classification and Similarity Models for Product Matching, in: Proceedings of NAACL-HLT 2018, New Orleans, Louisiana, June 1 - 6, 2018, Association for Computational Linguistics-ACL 2018, pp. 8–15. URL: https://aclanthology.org/N18-3002/
[43] N. Park, A. Kan, X. L. Dong, T. Zhao, C. Faloutsos, Estimating Node Importance in Knowledge Graphs Using Graph Neural Networks, in: Proceedings of the 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '19), August 4–8, 2019, Anchorage, AK, USA. ACM, New York, NY, USA, 2019. doi: 10.1145/3292500.3330855
[44] C. Alexandris, Evaluating Cognitive Bias in Two-Party and Multi-Party Spoken Interactions, in: Proceedings of Interpretable AI for Well-being: Understanding Cognitive Bias and Social Embeddedness (IAW 2019) in conjunction with AAAI Spring Symposium (SS-19-03), Stanford University, Palo Alto, CA, 2019. URL: http://ceur-ws.org/Vol-2448