Doctoral School of Philology

Faculty of Humanities

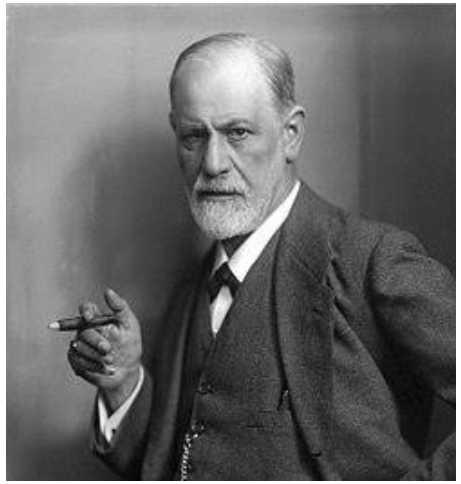
Moscow 2022

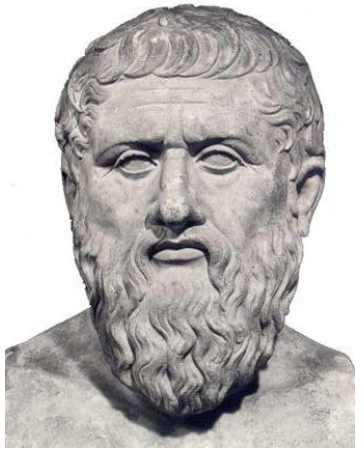
The Semantic Shift in Jokes: first computational experiments

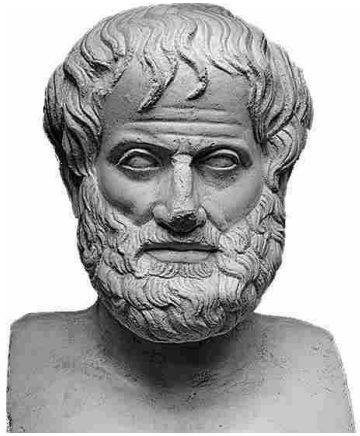
Author: Zakovorotnaia E. M.

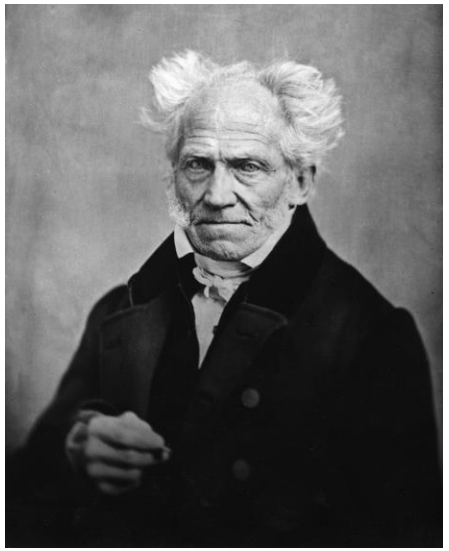
Humor research

- Philosophy (Plato, Aristotle, Cicero, I. Kant, A. Schopenhauer);
- Psychology (Z. Freud, T. Hobbes, F. Patterson, A. Behn, G. Spencer, V. Wundt, G. Dorfles);
- Linguistics (V. Raskin, S. Attardo, T. C. Veatch, A. Greimas, T. Todorov);
- Philology (V. Y. Propp, M. M. Bakhtin);
- Neuroscience (T. Deacon, S. Kennison).









Theoretical methodologies

- the Superiority Theory
- the Semantic Script Based Theory of Humor
- the Incongruity Theory
- the Relief Theory
- the General Theory of Verbal Humor

Applied automatic approaches

- Classification
- Generation
- Corpus Making

Research relevance

- lack of definition of the concept of humor;
- lack of objective criteria and characteristics of jokes;
- lack of an automatic system for generating short humorous texts with indicators above 0.7.

Definition of the semantic shift

The concept was applied to verbal humor in the Semantic Script Based Theory of Humor (SSTH) by linguist V. Raskin. It describes the formal representation of a verbal joke which main difference from the referential one is the language play.

All plots, ideas, characters, notions, events are presented by a frame or a script that is a certain knowledge structure corresponding to the meaning of a word or group words. The point is that scripts can be expressed by an lexically ambiguous word or expression. If two frames overlap in a joke, then the ambiguity of interpretations becomes visible and leads to the humorous effect.

V. Raskin claims that these scripts have to stay in a semantic opposition to each other and also be compatible with the same text.

Division of Jokes

Setup:

- everything else except punchline;

Punchline:

- last sentence of a joke;

- final part of the text sequence, if it's a one-shot liner;

Example: "*Is it true that cannibals do not eat clowns because they taste funny?*"
setup punchline

The corpus, code and resulting tables are presented in this repository on Github
https://github.com/zijane/semantic_shift_in_jokes

First Experiment

The first experiment proceeds with the word level. 128 out of 1000 jokes in our corpus have repetitive words in the first and last sentences. The selected humorous texts are used as an input of a BERT model 'bert-base-uncased' which creates word embeddings based on its context. The cosine similarity was calculated between two different embeddings of the same word in different positions.

The results demonstrate that 90% of jokes do not express any semantic shift since the point of a joke is not based on the word repetition; thus, understanding of a joke needs more precise work with contexts. There are about 3% of cases where a repeated word is associated with semantic shift.

Example: *"bus station is where a bus stops. A train station is where a train stops. On my desk i have a work station".*

In the rest of compound jokes, the frames' incongruity is tied to the other repetitive phrases or synonyms.

The corpus, code and resulting tables are presented in this repository on Github https://github.com/zijane/semantic_shift_in_jokes

Second Experiment

In the second experiment a working pipeline is completed with resolving of coreference – a type of textual or syntactic coherence in which two or more nominal groups describe the same object.

The cosine similarity is measured separately for each concept: the antecedent vector is compared to all embeddings of its anaphores. The extraction of coreferences increased the amount of corpus to 4328 lines; however, the number of jokes in the final corpus was 736.

We especially tagged these cases where the “antecedent-anaphora” pair is related to the semantic shift and how strong this connection is. We found out four cases:

- direct correlation with at least one “antecedent-anaphora” pair entirely reflects the semantic shift (6 %);
- indirect correlation with at least one “antecedent-anaphora” coreference is partially connected with the semantic shift, however, there is still not enough word or expression for the direct correlation (42%);
- absence of correlation in case of no “antecedent-anaphora” pairs with reflected semantic shift (50%);
- coreference extracted incorrectly in case of “antecedent-anaphora” pairs are composed incorrectly or the antecedent has no anaphores (2%).

The corpus, code and resulting tables are presented in this repository on Github https://github.com/zijane/semantic_shift_in_jokes

Second Experiment

Direct correlation with at least one “antecedent-anaphora” pair entirely reflects the semantic shift (6 %):

- Example one: "A man walking down the street meets a friend who has **a lobster** tucked under his arm. "Are you taking **that lobster** home to dinner ?" he asks . "No" , says friend , "**he** had his dinner and now i am taking **him** to the pictures."
- Example two: "Patient: doctor , doctor, **i** think **i** am a pair of curtains.

  Doctor: pull **yourself** together, **man**. "
- Example three: "Atoms were talking. Atom said to the **other**: "why are **you** crying ?" The **atom** replied: "**i** have lost an electron." The first atom said: "Are **you** sure ?". "Yes. " - replied the **other** - "**i** am positive!"

The corpus, code and resulting tables are presented in this repository on Github https://github.com/zijane/semantic_shift_in_jokes

Second Experiment

Indirect correlation with at least one “antecedent-anaphora” coreference is partially connected with the semantic shift, however, there is still not enough word or expression for the direct correlation (42%):

- Example one: "**A man driving on a highway** is pulled over by a police officer. The officer asks: "did **you** know **your** wife and children fell out of **your** car a kilometre back ?" A smile creeps onto the **man** face and **he** exclaims: "thank god ! thought **i** was going deaf ! "
- Example two: "**A man** walks into a restaurant and growls at the maitre d ' hotel. "Do you serve crabs here ?". The maitre d ' hotel responds: "We serve anyone. Have a seat, **sir** ."
- Example three: "learn the **rules** so you know how to break **them** properly."

The corpus, code and resulting tables are presented in this repository on Github https://github.com/zijane/semantic_shift_in_jokes

Second Experiment

Absence of correlation in case of no “antecedent-anaphora” pairs with reflected semantic shift (50%):

- Example one: "Did you hear about the rooster who stayed awake all night so that he could see where the sun went ? it finally dawned on him. "
- Example two: "what do an eagle and a lion have in common ? they both have wings , except for the lion."
- Example three: "A man died and his wife phoned the newspaper to place an obituary. She called the obituary department and said: "This is what i want to print : Bernie is dead. The man at the newspaper said: " But for you are allowed to print six words." the woman answered: "ok . then print : bernie is dead . toyota for sale."

The corpus, code and resulting tables are presented in this repository on Github https://github.com/zijane/semantic_shift_in_jokes

Second Experiment

Absence of correlation in case of no “antecedent-anaphora” pairs with reflected semantic shift (50%):

- Example one: "A man walking down the street meets a friend who has **a lobster** tucked under his arm. "Are you taking **that lobster** home to dinner ?" he asks . "No" , says friend , "**he** had his dinner and now i am taking **him** to the pictures."
- Example two: "Patient: doctor , doctor, **i** think **i** am a pair of curtains.

  Doctor: pull **yourself** together, **man**. "
- Example three: "Atoms were talking. Atom said to the **other**: "why are **you** crying ?" The **atom** replied: "**i** have lost an electron." The first atom said: "Are **you** sure ?". "Yes. " - replied the **other** - "**i** am positive!"

The corpus, code and resulting tables are presented in this repository on Github https://github.com/zijane/semantic_shift_in_jokes

Conclusions

The first experiment showed a count of jokes with repetitive words in the corpus and a difference between a pair embeddings for each case. Furthermore, resolving of coreference increased percent of the cases detected the semantic shift on 3 points in comparison with the results of the first one.

Experiments also demonstrated the not only necessity of coreferent detection but also parsing since many anaphora pronouns were often automatically extracted without its head nouns. In addition, parsing allows determining a head and dependent words for each antecedent and its anaphores which increases the quality of context processing and completes the formal description of frames probably defining the desired semantic shift.

The corpus, code and resulting tables are presented in this repository on Github https://github.com/zijane/semantic_shift_in_jokes

References

- Delvin et al., 2019 — J. Devlin, M.-W.Chang, K. Lee, K. Toutanova, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. URL: https://arxiv.org/abs/1810.04805
- Gabora, 2016. — L. Gabora, K. Kitto. Towards a Quantum Theory of Humour // Frontiers in Physics (section: Interdisciplinary Physics). URL: https://www.researchgate.net/publication/311843405_Towards_a_Quantum_Theory_of_Humour
- He He, Nanyun Peng, and Percy Liang. 2019. Pun Generation with Surprise. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 1734–1744, Minneapolis, Minnesota. Association for Computational Linguistics
- Morreall. 2020. — J. Morreall, “Philosophy of Humor", The Stanford Encyclopedia of Philosophy (Fall 2020 Edition), Edward N. Zalta (ed.), URL: https://plato.stanford.edu/archives/fall2020/entries/humor/.
- Paducheva E.V.2008. Utterance and its correlation with reality. Ed. 5, ispr. M.: URSS.
- Raskin, Attardo. 1994. — V. Raskin, S. Attardo. Non-literalness and non-bona-fîde in language: An approach to formal and computational treatments of humor. Pragmatics & Cognition, 2(1):31–69.
- Raskin, Attardo. 1991. — V. Raskin, S. Attardo. Script theory revis(it)ed: joke similarity and joke representation model. // Humor - International Journal of Humor Research, 4(3–4):293–348.
- Raskin. 1984. — V. Raskin. Semantic Mechanisms of Humor. Springer Netherlands, Dordrecht.
- Wiseman — R. Wiseman. 1001 jokes. URL: https://www.studocu.com/row/document/thammasat-university/new-media-studies/quirkology-1001-jokes-by-richard-wiseman/20822728
- Salvatore A. 1994. Linguistic Theories of Humor. Mouton de Gruyter, New York. p. 335.

The corpus, code and resulting tables are presented in this repository on Github https://github.com/zijane/semantic_shift_in_jokes

Thank you for attention!

The corpus, code and resulting tables are presented in this repository on Github https://github.com/zijane/semantic_shift_in_jokes