Cogn Process
DOI 10.1007/s10339-008-0221-3

LETTER TO THE EDITOR

GraPHIA: a computational model for identifying phonological jokes

Narayanan Srinivasan · Vani Pariyadath

Received: 10 January 2008 / Revised: 9 June 2008 / Accepted: 20 June 2008

Abstract Currently in humor research, there exists a dearth of computational models for humor perception. The existing theories are not quantifiable and efforts need to be made to quantify the models and incorporate neuropsychological findings in humor research. We propose a new computational model (GraPHIA) for perceiving phonological jokes or puns. GraPHIA consists of a semantic network and a phonological network where words are represented by nodes in both the networks. Novel features based on graph theoretical concepts are proposed and computed for the identification of homophonic jokes. The data set for evaluating the model consisted of homophonic puns, normal sentences, and ambiguous nonsense sentences. The classification results show that the feature values result in successful identification of phonological jokes and ambiguous nonsense sentences suggesting that the proposed model is a plausible model for humor perception. Further work is needed to extend the model for identification of other types of phonological jokes.

Keywords Humor · Puns · Computational model · Graph theory

Introduction

Humor perception research in cognitive science is not well advanced and this may be attributed partly to the difficult, but not impossible, task of studying a higher order cognitive function like humor that is dependent on processes including language and emotions. Several behavioral experiments and neuroimaging studies have been performed to study humor perception (Vaid et al. 2003; Ruch et al. 1993; Shammi and Stuss 1999; Goel and Dolan 2001; Moran et al. 2004). However, computational research on humor perception is scarce.

Computational humor research consists of research on both humor generation and humor perception. Several models have been proposed for humor generation with varying amounts of success (Attardo and Raskin 1994; Binsted 1996; McDonough 2001; Stock and Strapparava 2002). At the same time, there is relatively less work on models for humor perception. Some models or algorithms proposed for humor perception or identification are the Knock Knock Joke Recognizer (Taylor and Mazlack 2004), Katz's neural model (Katz 1996) and Japanese Pun Analyzer (Yokogawa 2002). General theories for humor perception (Mulder and Nijholt 2002) share a common limitation in not being quantifiable (Srinivasan and Pariyadath 2008). We believe that existing theories can be improved by building better computational models.

All theories of humor perception generally agree on the necessity for identifying an incongruity or violation to perceive humor. In fact, humor perception may be considered to be the resolution of such an incongruity (Ritchie 1999). The violation theory (Veatch 1998) states that there are three necessary and sufficient conditions for humor perception: violation (V) of a moral commitment that the perceiver holds important, normalcy (N) perceived in the situation, and simultaneity of both and N and V. Taking inspiration from violation theory, this paper proposes a graphical phonological humor identification algorithm—GraPHIA. Such a specialized model solely for phonological jokes or puns has been constructed for several reasons. Recent neuroimaging studies show that phonological jokes

N. Srinivasan (✉) · V. Pariyadath
Centre for Behavioural and Cognitive Sciences,
University of Allahabad, Allahabad 211-002, India
e-mail: nsrini@cbcs.ac.in

activated different areas compared to semantic jokes (Goel and Dolan 2001). While puns showed activations in the speech production regions centered in the left hemisphere, semantic jokes showed activations in bilateral temporal lobe networks. Thus, we have reason to believe some processing is strictly distinct and separable for phonological jokes. Modeling identification of puns has the added benefit of studying a humor subtype, seemingly devoid of an affectual component. Puns have traditionally been considered a weak humor subtype (Veatch 1998) perhaps due to this questionably minimal affect. Phonological jokes are generally classified into three categories: homophonic, homographic and double-sounded jokes. In this paper, we will focus on homophonic puns. Homophonic jokes contain wordplay involving homophones or words that sound the same but are spelled differently. With puns, the violation lies primarily in the wordplay, where one word meaning triggers a violating interpretation. In the case of homophonic puns, the similar sounding word (tears in J1) causes the resolution, as in:

J1: *It was an emotional wedding. Even the cake was in tiers* (pun of the day)

GraPHIA makes use of graph theoretical ideas to compute new features that may correspond to the detection and resolution of violation to identify input as normal sentences, phonological jokes or ambiguous nonsense sentences.

Methodology

The model assumes that the language comprehension processes have already carried out relevant semantic and phonological processing of the words in the joke input. In this regard, the proposed model does not claim to be carrying out any specific language processing and assumes that semantic and phonological information is available for further humor-related computations. Therefore, the input to GraPHIA is in the form of a set of semantic and phonological weights between the nodes, which are assumed to be available from an earlier stage of processing. Such an assumption would further imply that at the time of hearing or reading a joke, only the relevant words (or rather concepts) are present in working memory. On encountering a violation in the text associated concepts are brought into working memory in an attempt to resolve the violation. In the case of puns, the associated concepts would be homophones of words in the input. Any possible joke may be considered to be processed in two phases. Initially, the semantic congruity of the text is tested and if faced with a violation, relevant homophones are considered in an attempt to resolve this violation. If this resolution is successfully carried out, a pun is identified.

Our model assumes two networks—a semantic network and a phonological network, which conforms to present findings from neuroimaging studies (Goel and Dolan 2001). In both networks, words are represented by nodes that are sparsely connected by edges with semantic and phonetic weights. In the semantic network, these weights range from very low for synonyms to very high for words that are semantically not related to each other. In the phonological network, the weights range from very low for homophones to very high for no acoustic or phonological similarity. When the joke concepts are brought into working memory, a fully connected temporary sub-graph (based on the large semantic and phonological network stored in the long-term memory) containing only these concepts is constructed. An example is shown in Fig. 1 for joke J2.

J2: *Atheism is a non-prophet organization*

Representing the words as nodes as part of a network (graph) enables us to apply graph theory to compute new features for identification of puns. The input to GraPHIA is in the form of an adjacency matrix $\boldsymbol{A}$ where each entry $\boldsymbol{A}_{ij}$ corresponds to the weight of the connection between the nodes i and j. The algorithm in effect makes computations with two types of weights—semantic as well as phonological. To estimate violation in the joke text, we compute the strength of the minimal spanning tree (MST) α from the adjacency matrix $\boldsymbol{A}$ for the semantic subgraph consisting of n nodes. Among all the spanning trees of a weighted and connected graph, the one with the least total weight is called an MST (Deo 1974). If α/n is between the two thresholds, then the input text is treated as a potential joke. If α/n is greater than the upper limit the input is treated as an ambiguous nonsense sentence with no acceptable meaning (as in A1) and if the α/n value is below the lower limit the input text is identified as a non-joke. If the joke input is found to contain a violation, then all relevant homophones are taken into consideration. If no homophones are found, the text is identified as a non-pun. If homophones are found, the MST is recomputed for the

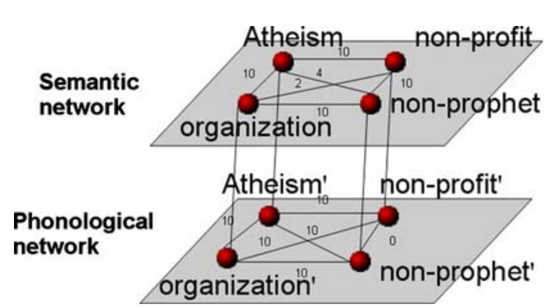

Fig. 1 Graphical structure for semantic and phonological networks for Joke J3

hybrid matrix (***H***). The hybrid matrix is obtained from the semantic (***S***) and phonological (***P***) weight matrices where each element $\boldsymbol{H}_{ij} = \text{Min}\ (\boldsymbol{P}_{ij}, \boldsymbol{S}_{ij})$. The strength of the MST for the hybrid matrix β is compared with α by computing a discriminant to determine the presence of a pun. The discriminant δ is computed from

$$\delta = \frac{(\alpha - \beta)}{n} \tag{1}$$

where n is the number of content words in the joke input. This discriminant value will depend on whether a resolution is possible with either one of the homophones. If δ is above the threshold value, the input is treated as an ambiguous nonsense sentence with no meaning as in example A1. While this indicates the presence of wordplay, it does not establish whether the wordplay combines two equally valid interpretations of the homophone involved. To evaluate this condition, we compute the MST α' of the semantic subgraph containing the joke input with the homophone replacing its counterpart, and now compute equivalence γ:

$$\gamma = \alpha - \alpha' \tag{2}$$

The equivalence value is used to determine whether the given text is a joke or ambiguous nonsense sentence with one acceptable meaning as in example A2.

A1: *When they found the key a kettle bloomed*
A2: *She ate my breakfast role*

The computations performed by GraPHIA are described below:

(a) *Compute the strength of MST α from the adjacency matrix of the semantic network of the given input text.*
(b) *If a violation/incongruity or a high value of (α/n) is found and α/n is also less than a threshold*
   a. *Include all relevant homophones.*
   b. *If no relevant homophones found, no pun identified.*
   c. *Else Compute MST β for the hybrid adjacency matrix containing semantic and phonological weights*
      i. *Compute discriminant δ from α and β.*
      ii. *If δ is between the two thresholds*
         1. *Compute α' of the subgraph containing the joke input with the homophone replacing its counterpart.*
         2. *Compute (equivalence) $\gamma = \alpha - \alpha$'*
         3. *If γ is less than the threshold then a pun is identified.*
         4. *Else the input is identified as an ambiguous sentence with one acceptable meaning.*
      iii. *Else the input is identified as an ambiguous sentence with absolutely no meaning.*
(c) *If α/n is greater than the threshold the input is an ambiguous sentence with absolutely no meaning.*
(d) *Input is a normal sentence.*

The data set consisted of 40 homophonic puns, 40 normal sentences and 40 ambiguous nonsense sentences. Of the 40 ambiguous nonsense sentences, 20 were those in which ambiguity was present due to the presence of a homophone but there was no meaning to the sentence as in A1, and 20 were those in which the wrong homophone was present in the sentence as in A2. In GraPHIA, the MSTs were computed using Kruskal's Algorithm (Deo 1974). While presenting the input, only the content words in the text were included for constructing the subgraphs. For example, for joke J1, the content words were emotional, wedding, cake, and tiers. For joke J3 the content words were atheism, non-prophet and organization. The weights were chosen by the authors since there is at present no simple acceptable method to define a single value representing the semantic or phonological similarity between the two words. To classify the data set into normal sentences, phonological jokes, and ambiguous nonsense sentences, we computed an explicit threshold and evaluated the classification performance with such a specific threshold.

Results

The algorithm was applied to all the sentences in the data set and sample values of α, β, γ and δ for homophonic puns, ambiguous nonsense sentences and normal sentences (see Appendix for some sample input sentences) are given in Table 1. For example, the semantic and phonological subgraphs for the joke J2 are shown in Fig. 1. Similarly the subgraphs for a normal sentence N1 and ambiguous nonsense sentences A3 and A4 are shown in Figs. 2, 3 and 4.

N1: A bee stung her on the cheek
A3: The car's mane was very shrewd
A4: There was a flee on the dog's back

α is computed for the subgraph containing only one of the homophones, α' is computed for the subgraph containing the other homophone and β is computed for the subgraph containing both the homophones. At the first stage, violation (α/n) is computed and a minimum and maximum threshold of 2.5 and 5.0 were used for classification. If α/n was less than or equal to 2.5 then the sentence was classified as a normal sentence. If α/n was greater than or equal to 5.0 then the sentence was classified as an ambiguous nonsense sentence. Similarly in the next step, a threshold value of 0.5 for the discriminant (δ) value was used to identify ambiguous nonsense sentences of type A1. Finally, a threshold value of 5.0 was used to separate puns

Table 1 Feature values for selected homophonic puns (J), normal sentences (N) and ambiguous nonsense sentences (A1 and A2)

Joke	Type	$\alpha/n1$	δ	γ
Pisa	J	4.66667	1.33333	2
Criminal	J	4	2.66667	0
Dentist	J	4	2.66667	0
Gold	J	4.5	3	4
Atheism	J	4.66667	3.33333	0
Steel	A1	5.5	0	0
Key	A1	5	0	0
Car	A1	10	3.33333	0
Breakfast	A2	4	2.66667	8
Cruise	A2	4	2	8
Flee	A2	3.33333	2	6
Werewolves	N	1.5	0	0
Deer	N	2	0	0
Bee	N	2.5	0	0
Plate	N	1.5	0	0
Cabbage	N	2	0	0

from ambiguous nonsense sentences of type A2. The model could correctly identify all input sentences using the pre-assigned thresholds.[1]

Discussion

The proposed model based on graph theoretical ideas and relevant findings from cognitive neuroscience proposes three new features that could be computed from a small semantic network and phonological network for phonological joke identification. GraPHIA could accurately classify given input sentences as puns, normal sentences and ambiguous nonsense sentences. These results suggest that it could be considered to be a plausible model for perception of phonological jokes. Moreover, GraPHIA provides a way of computing quantifiable features that may be linked to the funniness of a phonological joke or response times to a phonological joke. The classification of an input sentence as a homophonic pun may depend on the phone that is used in the text. For example, the homophonic pun, "*He enjoyed being a chimney sweep because it suited him*" is not classified as a pun if the homophone "suited" is used (the violation value α/n is lower than the threshold) where as it is classified as a pun if "sooted" is used (the violation value α/n is greater than the threshold). These instances also provide a way of empirically testing the model with sentences whose funniness may depend on the homophone that is used in the text.

[1] It is possible that the computation of thresholds is data set specific and may not generalize very well to other data sets. Hence, we also used a two-layer feed forward neural network to classify the given text input into the relevant classes which does involve a fixed threshold. The network was trained with the Levenberg–Marquardt algorithm and the simulation was implemented in MATLAB. Half the data set used for training and the other half of the data set used for testing. An accuracy of 98.3% was obtained with the neural network classifier indicating that the novel features can be used to identify normal sentences, homophonic puns and the two types of nonsense ambiguous sentences even without using pre-assigned explicit thresholds.

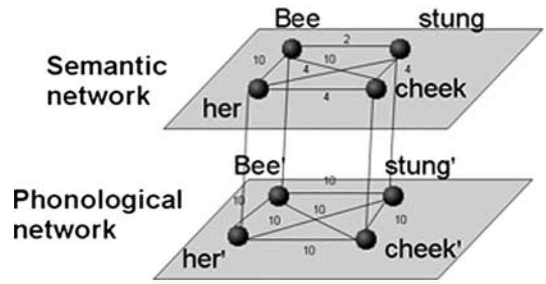

Fig. 2 Graph representation of the semantic network for N1

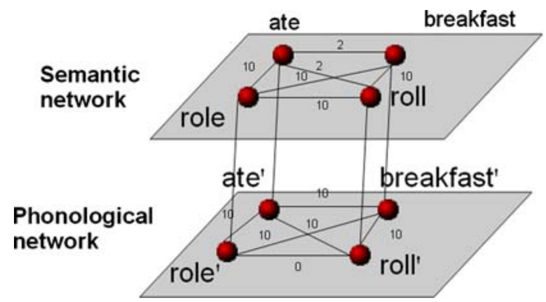

Fig. 3 Graph representation of the semantic network for A3

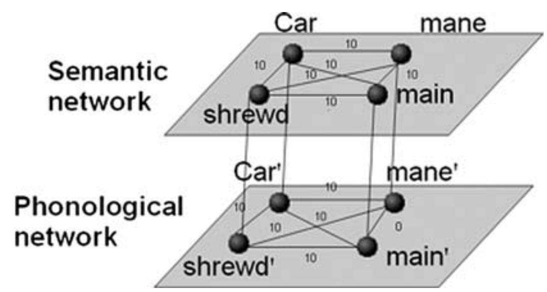

Fig. 4 Graph representation of the semantic network in A4

The formation of the two small subgraphs based on content words from the given sentences and the feature computation are hypothesized to occur in working memory. The assertion that such a *concurrent activation* of relevant concepts in working memory is required for humor processing is by no means unique (Attardo 1997; Nerlich and Clarke 2001; Vaid et al. 2003). The proposed feature computation may be performed in the prefrontal networks associated with working memory as suggested by neuroimaging findings (Shammi and Stuss 1999). In the proposed algorithm, resolution was attempted by bringing relevant homophones into working memory. If no resolution were possible, perhaps other related semantic concepts could be brought into working memory. In comparison with violation theory (Veatch 1998) that refers to N (normalcy) and V (violation) for perceiving humor, GraPHIA computes three values—α/n, δ and γ. The discriminant δ is dependent on the α and β values, which could be likened to measures of violation and normalcy, respectively. The use of the other variable, the equivalence γ, evaluates the way the homophone links two different interpretations of the scenario and may be likened to simultaneity in the violation theory.

How does GraPHIA compare with other computational models for humor perception? It is to be noted that most theories for humor perception including the violation theory are not quantifiable (Srinivasan and Pariyadath 2008). Focusing on computational models, the KK joke recognizer (Taylor and Mazlack 2004) uses N-grams to detect word play by generating possible wordplay sequences. The joke recognizer has many drawbacks with the most significant being the inability to tell whether a sentence is a joke or it simply contains wordplay. Not all wordplays are funny. Yokogawa's pun analyzer (Yokogawa 2002) uses articulation similarities to generate possible pun candidates in places of ungrammatical parts in sentences. Both the neural model by Katz (1996) and the nonlinear dynamics based model by Paulos (1980) have not been tested with any humorous sentences and have not been quantified to make any evaluations possible. GraPHIA performs much better than the KK joke recognizer and the pun analyzer strengthening its plausibility as a model for humor perception at least with respect to phonological jokes and also as an effective classification algorithm for identifying puns in natural language.

GraPHIA proposes novel computable features that may be correlated with funniness ratings and reaction times for responding to jokes, which can be tested experimentally. An important issue is the lack of detailed language processing performed by the model. While other computational models for humor perception or generation have incorporated some module for wordplay detection, the model described by this paper skips such a step. GraPHIA assumes that language comprehension is successfully carried out and such successful language processing is necessary for perception of verbal humor as well as comprehending normal sentences. The model solely focuses on additional processing (presumably occurring in working memory) needed for perceiving the puns. This essentially means that the algorithm can be combined with other language-processing models. In addition, the weights used in the input can be improved by incorporating weights from future semantic and phonological networks that may be developed with further progress in language research.

In conclusion, GraPHIA proposes a novel graph theory-based algorithm that successfully classifies homophonic jokes and presents a putative model for phonological humor perception. Further work is required to validate the model with experimental results and to elaborate on the role of computations in working memory in perceiving humor.

Appendix: List of sentences for which feature values are shown in Table 1

List of jokes

[Pisa] Italian building inspectors in Pisa are leanient.
[Criminal] A criminal's best asset is his lie ability.
[Dentist] Be kind to your dentist because he has fillings too.
[Gold] You can make gold soup by putting in 24 carrots.
[Atheism] Atheism is a non-prophet organization.

List of ambiguous sentences

[Car] The car's mane was very shrewd.
[Key] When they found the key a kettle bloomed.
[Steel] A blue steel waited calmy.
[Breakfast] She ate my breakfast role.
[Cruise] The cruise finished construction on the building.
[Flee] There was a flee on the dog's back.

List of normal meaningful sentences

[Werewolf] Werewolves are fictitious creatures like elves.
[Deer] Not all animals are as agile and graceful as a deer.
[Bee] A bee stung her on the cheek.
[Plate] She washed the plates with soap and water.
[Cabbage] Cabbages are like lettuce because they are both vegetables.

References

Attardo S (1997) The semantic foundations of cognitive theories of humor. Humor Int J Humor Res 10:395–420

Attardo S, Raskin V (1994) Non-literalness and non-bona-fide in language: an approach to formal and computational treatments of humor. Pragm Cogn 2:31–69

Binsted K (1996) Machine humor: an implemented model of puns Doctoral Dissertation. University of Edinburgh, Edinburgh, Scotland

Deo N (1974) Graph theory with applications to engineering and computer science. Prentice-Hall, India

Goel V, Dolan RJ (2001) The functional anatomy of humor: segregating cognitive and affective components. Nat Neurosci 4:237–238. doi:10.1038/85076

Katz B (1996) A neural invariant of humour. Proc Internatl Workshop Comp Humor University of Twente, Netherlands, pp 103–109

McDonough CJ (2001) Mnemonic string generator: software to aid memory of random passwords. Technical Report. CERIAS, Purdue University

Moran JM, Wig GS, Adams RB Jr, Janata P, Kelley WM (2004) Neural correlates of humor detection and appreciation. Neuroimage 3:1055–1060. doi:10.1016/j.neuroimage.2003.10.017

Mulder MP, Nijholt A (2002) Humour research: state of the art. Technical Report No. 02-34, University of Twente, Netherlands

Nerlich B, Clarke D (2001) Ambiguities we live by: towards a pragmatics of polysemy. J Pragmatics 33:1–20. doi:10.1016/S0378-2166(99)00132-0

Paulos J (1980) Mathematics and humor. University of Chicago Press, Pun of the day. http://www.punoftheday.com

Ritchie G (1999) Developing the incongruity-resolution theory. Proc AISB '99 Symp creative language: humor and stories, Edinburgh, pp 78–85

Ruch W, Attardo S, Raskin V (1993) Towards an empirical verification of the general theory of verbal humor. Humor Int J Humor Res 6:123–136

Shammi P, Stuss D (1999) Humor appreciation: a role of the right frontal lobe. Brain 122:657–666. doi:10.1093/brain/122.4.657

Srinivasan N, Pariyadath V (2008) Dissecting the frog: computational approaches to humor perception. In: Srinivasa N, Gupta AK, Pandey J (eds) Advances in cognitive science, vol 1. Sage Publications, New Delhi, pp 199–215

Stock O, Strapparava C (2002) Humorous agent for humorous acronyms: the HAHAcronym project. Humor Int J Humor Res 16:297–314. doi:10.1515/humr.2003.016

Taylor J, Mazlack LJ (2004) Computationally recognizing wordplay in jokes. Proc Cog Sci Soc, Chicago, pp 1315–1320

Vaid J, Hull R, Heredia R, Gerkens D, Martinez F (2003) Getting a joke: the time course of meaning activation in verbal humor. J Pragmatics 35:1431–1449. doi:10.1016/S0378-2166(02)00184-4

Veatch TC (1998) A theory of humor. Humor Int J Humor Res 11:161–215

Yokogawa T (2002) Japanese pun analyzer using pun articulation similarities. Proc IEEE Intern Conf Fuzzy Sys, Honolulu, USA, pp 1114–1119