INFORMATION RETRIEVAL IN THE SERVICE OF GENERATING NARRATIVE EXPLANATION
What we Want from GALLURA

Ephraim Nissan[1] and Yaakov HaCohen-Kerner[2]
[1]*Dept. of Computing, Goldsmiths' College, Univ. of London, 25–27 St. James, New Cross, SE14 6NW, London, U.K.*
[2]*Department of Computer Science, Jerusalem College of Technology (Machon Lev), P.O.B. 16031, Jerusalem, Israel*

Keywords: Information extraction, Explanation generation, Story generation.

Abstract: Information retrieval (IR) and, all the more so, knowledge discovery (KD), do not exist in isolation: it is necessary to consider the architectural context in which they are invoked in order to fulfil given kinds of tasks. This paper discusses a retrieval-intensive context of use, whose intended output is the generation of narrative explanations in a non-bona-fide, entertainment mode subject to heavy intertextuality and strictly constrained by culture-bound poetic conventions. The GALLURA project, now in the design phase, has a multiagent architecture whose modules thoroughly require IR in order to solve specialist subtasks. By their very nature, such subtasks are best subserved by efficient IR as well as mining capabilities within large textual corpora, or networks of signifiers and lexical concepts, as well as databases of narrative themes, motifs and tale types. The state of the art in AI, NLP, story-generation, computational humour, along with IR and KD, as well as the lessons of the DARSHAN project in a domain closely related to GALLURA's, make the latter's goals feasible in principle.

1 CONCEPTUAL & TECHNICAL BACKGROUND

In the history of full-text IR, tools for retrieval from very large historical corpora in Hebrew and Aramaic were prominent, with the RESPONSA project (see e.g. Choueka, 1989a, 1989b; Choueka et al. 1971, 1987). Before the rise of Web search engines, RESPONSA tools were the ones which achieved the more far-reaching effects on society, because how they empowered the retrieval of legal precedents in rabbinic jurisprudence, thus affecting especially legal practice of family law in Israel (as for family law, in the Ottoman successor states, the usual jurisdiction is the courts of the various religious communities).

Religious cultures, as being the "consumers" of religious texts, were, in a sense, the customers of a considerable portion of early projects in IR: apart from RESPONSA, whose corpora comprise the Jewish texts from the sacred sphere through the ages, this was also the case of Padre Busa's *Index Thomisticus* in Milan, and of the humanities computing at the Abbey of Maredsous, in Belgium.

Exegesis (such as biblical interpretations) and homiletics involve layers of texts, where a secondary text refers to and either just quotes, or discusses, some locus in the primary text; or then (as in the Jewish *aggadic midrash*) expands on a biblical narratives, filling the gaps where the primary text is silent. Collections of *aggadic midrash* from late antiquity (e.g., the *Midrash Rabbah*) or the Middle Ages (e.g., *Yalqut Shim'oni*) are a digest of a multitude of homilies on biblical fragments of texts, developing several often alternative ideas and subnarratives. Cf. Hirshman (2006), Braude (1982), Fishbane (1993), Hartman and Budick (1986).

* *HyperJoseph* is a hypertextual tool on the story of Joseph in *Genesis,* with the secondary texts elaborating on it (Nissan and Weiss, 1994).
* DARSHAN is a tool that invents homilies in Hebrew (HaCohen-Kerner et al. 2007). Retrieval in DARSHAN is intensive, and so is the use of networks of lexical concepts.

DARSHAN generates ranked sets of either one-sentence or one-paragraph homilies. While producing its output, DARSHAN is able to quote

Nissan E. and HaCohen-Kerner Y..
INFORMATION RETRIEVAL IN THE SERVICE OF GENERATING NARRATIVE EXPLANATION - What we Want from GALLURA.
DOI: 10.5220/0003688304790484
In *Proceedings of the International Conference on Knowledge Discovery and Information Retrieval* (KDIR-2011), pages 479-484
ISBN: 978-989-8425-79-9

from Scripture, to search for an occurrence elsewhere in the textual canon, to replace words or letters, to resort to puns, to interpret a word as an acronym, and so forth. Use is made of patterns which consist of canned text with places where to plug in strings obtained through IR and manipulation. The user supplies as input a biblical verse, or a sentence, or a set of words, and also specifies which devices should be applied. Filters applied to the candidate output are alert, e.g., to positive vs. negative connotation.

The quality of an individual output homily is assessed as a sum of weighted factors, including: length (as an indicator of complicacy); the percentage of relevant words in the homily, out of the total of words in the homily how many sentences there are; how complex it was to insert every motif into the homily generated; how many motifs were actualized in the output homily being evaluated; how many transformations were carried out; how many words were replaced in the homily.

Having mentioned acronyms, consider that HaCohen-Kerner et al. (2010b) discussed an abbreviation disambiguation system for rabbinic texts in Hebrew or Aramaic. Cf. Stock and Strapparava (2005) on the HAHA project, whose purpose is the humorous interpretation of acronyms. As to connotations, Strapparava and Valitutti (2004) described an affective extension of WordNet.

2 FUNCTIONS IN GALLURA

The GALLURA project seeks to develop software that would interpret in Hebrew names by folk-etymology, but in the context of a generated narrative (aetiological tales, usually brief or even very brief). The most closely studied model is a large textual corpus of playfully creative writing that embodies midrashic literary devices, by explaining fancifully place-names of names for animal kinds.

The GALLURA project, now in the design phase, requires, among the other things, capabilities of *story-generation,* and of generating a *playful explanation.* By themselves, these two tasks draw upon three areas in AI:

- *explanation synthesis* (for which, see e.g. Schank, 1986, 1994; Walton, 2004),
- *story-generation* (see e.g. Liu and Singh, 2002; Lönneker et al., 2005; and a long survey in Nissan, 2011a: Ch. 5), and
- *computational humour* (see e.g. Stock et al., 2002; Ritchie, 2004; Waller et al., 2009). Humour studies are interdisciplinary.

Moreover, GALLURA needs skills from *computational linguistics,* including some that thus far were modelled by linguistics, but not computationally:

- *folk-etymology* (see e.g. Kirwin, 1985; Coates, 1994; Baldinger, 1973; Zuckermann, 2006), and
- *phono-semantic matching* (*PSM*), a discussion of which is found in Zuckermann (2000, 2006).

For example, one of several PSM rules as occurring in neologisation by adapting a foreign term (Zuckermann 2000) is as follows (where SL is the source language. TL is the target language):

SL y 'b' → $TL_{(+PSM)}$ x 'b' ← TL x 'a'

x is phonetically similar to y; a is similar to b

That is to say, the PSM introduced a new sense: this was a PSM produced by shifting the meaning of a pre-existent word in the *target-language* (*TL*). Another rule of *camouflaged borrowing* (*ibid*.) is:

SL y 'b'→$TL_{(+PSM)}$ {x}+{z} 'b'←TL {x} 'a', {z}

x is a lexical morpheme (e.g. root) that is phonetically similar to y;
z is a grammatical morpheme (e.g. noun-pattern);
{x}+{z} is one word; a is similar to b

GALLURA should also have quality evaluation capabilities, e.g., evaluating a story generated (Peinado and Gervás, 2006), or evaluating morality within a story (Reeves, 1991). We also need to resort to *computational argumentation:* some such current research into argumentation in computer science looks into legal narratives (Bex, 2011).

Explanation as sought in GALLURA need not necessarily be realistic; it is non-bona-fide (like in humour), and must conform to a set of conventions, of which realism is just a particular case (cf. Nissan, 2008). There are constraints on style: the output text generated conforms to the early rabbinic linguistic stratum and style (thus emulating the *aggadic midrash*), with constraints on which lexical items or morphological forms can be selected.

Rabbinic stylemes are the subject of current IR research, including in the CUISINE text classifier. So are the identification of rabbinic citations, and chronological classification based on them. In fact, HaCohen-Kerner et al. (2010a) discussed stylistic feature sets for classification in CUISINE. Automated identification of citations from rabbinic texts has been researched (HaCohen-Kerner et al., 2010c). Automated classification of rabbinic

responsa by period based on what they cite or are cited by, was attempted successfully: HaCohen-Kerner and Mughaz (2010) defined and effectively applied "various kinds of 'iron-clad', heuristic and greedy constraints defining the birth and death years of an author based on citations referring to him or mentioned by him."

3 THE MULTIAGENT ARCHITECTURE

Several capabilities are required of GALLURA, and many of them require retrieval. Fig. 1 shows **Coalition1** of agents, i.e. agents that often interact among themselves. The control sequence is opportunistic, according to the needs of the various agents while they tackle a (sub) problem during a particular run. They broadcast their need for help to the other agents, and contract out the task. Some agents however interact in a privileged manner with one or more other gents, as they for a "coalition".

Both the syntax agent and the stylemic agent have to emulate early rabbinic language, but the pool of stylemes and more abstract modes comprising stylemes need actually be wider. Fig. 2 shows the interplay of other coalitions of agents. In the Lexicon, expected associations or behaviour are triggered through *demons,* procedural code activated upon access to individual lexical entries.

Coalition5 comprises an Encyclopedic agent, and a Commonsense agent. The latter comprises two modules: Concept-centred commonsense, and Situational commonsense. Both the Emplotment agent, and the Tex-generation agent closely interact with the Argumentation agent.

4 A SIMPLE EXAMPLE: *AQUA* & *GENESIS* 1:9

It is usually proper nouns that are playfully etymologised in the modern, archaising Hebrew narrative corpus which is the main model for GALLURA, and whose own model is the already mentioned early rabbinic genre of the *aggadic midrash.* Nevertheless, sometimes common nouns are folk-etymologised as well, and most often these are non-Hebrew words.

Here is a concise example. The input is Latin *aqua* 'water'. In the model corpus, there is this item:

> *Ma ra'ú Bnei Rómi, še-hém qorín et ha-máyim 'aqwa (aqua)? Le-fí še-katúv: "yiqqawú ha-máyim".*

Here is a translation of this Hebrew text:

> Why [literally: what did they see], the Romans [lit.: The Sons of Rome], that they call water *aqua?* Because [lit.: to mouth of] it is written [in Scripture]: "Let the water be gathered".

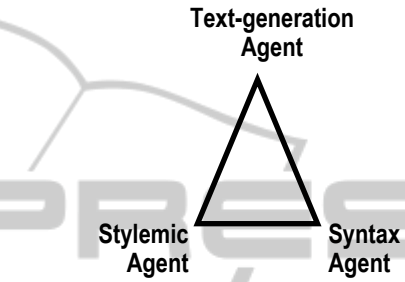

Figure 1: **Coalition1** of agents.

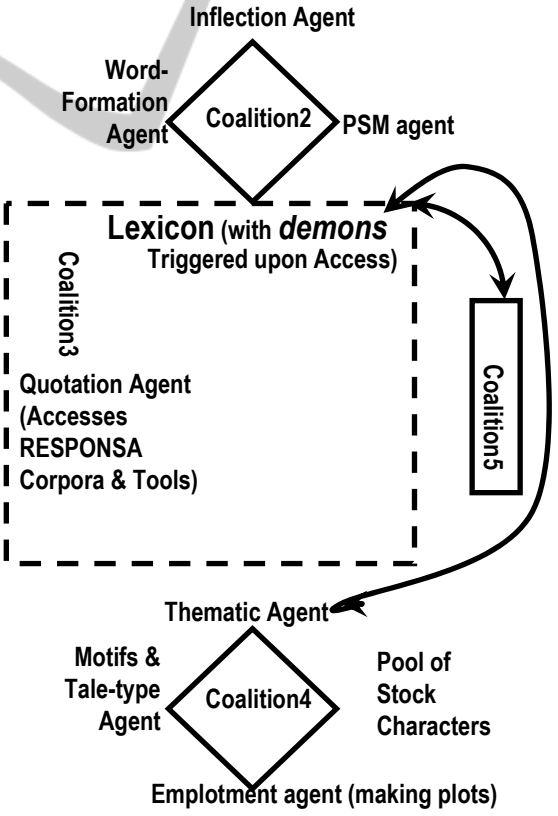

Figure 2: The interplay of coalitions of agents.

In fact, the intertextual reference is to *Genesis* 1:9. The verbal form *yiqqawú* (passive future, 3rd

person plural) is from the root *qwh.* Corradicals one can find in the Hebrew Bible include the verb and noun for fluids gathering, for hoping and hope (the word for 'hope' also has the little known sense 'string'), and the noun now used for 'line'. Etymologically unrelated, *Qwe* also occurs, being the name of a horse-trading land in Anatolia with whom and with Egypt King Solomon traded in such animals. Finding the apparent corradicals is trivial, using the IR and NLP tools of the RESPONSA project. What does require AI instead is for software to be able to notice that *Genesis* 1:9, because it is about water (and during an act of creation), is splendidly apt an occurrence of the input *aqua,* which PSM spuriously proposes as a derivative of the root *qwh* (Semitic roots are "triliteral").

There are features of the example considered, pertaining to the lexicon, morphology, and style, which clearly belong to the Mishnaic (i.e., early rabbinic) historical stratum of Hebrew. Beginning with a question, and in particular with one of the many ways of asking 'Why' in Hebrew (i.e., lit. "What did they see?"), which involve **Coalition1** and **Coalition2,** the Lexicon.

Asking and answering here also involves some rather rudimentary involvement of argumentation. A shortcut would be to use a canned-text encoding of a pattern, in the manner of DARSHAN. Actually however there is some sophistication in the example considered, because we are not abstractly taking about Latin; rather, the expression is made concrete, with the Sons of Rome being invoked from the Pool of stock characters. This dovetails with the underscoring of their agency, when the option selected for saying 'Why' is "What did they see?"

5 A COMPLEX EXAMPLE: BABEL TO LAOS

The following would be a much more difficult example for GALLURA to replicate, and both retrieval and manipulation would be intensive and laborious. In the model corpus we use, place-names around the world are explained by both playful etymology, and fantasy history narratives. It is often the case that a story is told about one of the human groups leaving the Tower of Babel. The Generation of the Division (*Dór ha-Pallagá*) or the Ones Leaving the Tower (*Yots'ei ha-Migdal*) would be often resorted to in GALLURA's Pool of stock characters. Let us consider a story on Laos.

"Teach us, Sir" (*yelammédenu Mar,* a cliche especially associated with the lost rabbinic *Midrash Yelammedenu*), "What did the Nations see" (i.e., 'why': *má ra'ú ha-'ummót*), "that they call" (*še-qorín:* a Mishnaic verbal inflection) "one of them Laos" (*achát mehén Lá'os*). "I shall answer you immediately!" (*Af aní mešivkhem mi-yád!* a cliché). A ready pattern of argumentation: "Instead of [lit.: Until] you asking why that nation is called Laos" (*'Ád še-attém šo'alín lámma otáh 'ummá qruyá Lá'os*), "be asking what did the Sons of Greece see" (*hevú šo'alín ma ra'ú Bnei Yaván*), "that all populations" (*še-kól 'okhlosín,* itself a Green loanword in Hebrew) "were called in their mouths [i.e., by them] λαός" (*niqre'ú be-fihém lá'os*). "Once the Ones Leaving the Tower went out of Babel" (*Keván še-yats'ú Yots'éi ha-Migdál mi-Bavél*), "they were tired (*le'ín*) and exhausted on the road" (*hayú le'ín u-me'uyyafín ba-dárekh*).

Sustained walking is tiresome, and one term for 'tired' is related by PSM to *Laos.* Now, consider that in a crowd (a spawned demon would inform GALLURA), you would expect somebody trying to sell snacks and drinks, unless circumstances exclude this (e.g., if it's a day of fast, or a famine causes starvation). Such a situational cliché is funny if it does not quite match the situation at hand. The theme of the exodus from Babel, in the model corpus, often has a wise old man advise the crowd, but some other time, some individual takes advantage, being cunning rather than altruistic.

"The more astute among them" (*'Armumiyyín se-bahém*), "who were traders and vendors of edibles" (*še-hayú ba'aléi praqmátya* [a typical early rabbinic term] *u-mokhréi mezonót*), "this way they were speaking to them" (*kákha hayú 'omrím lahém*): "Let the legs be strong!" (*Techezáqna ha-ragláyim!*). The latter contains a Biblical Hebrew verbal form, the 3rd person plural feminine (as 'legs' are feminine in Hebrew), whereas Mishnaic Hebrew discarded that form, using the masculine. As this is a modified quotation, using a Biblical Hebrew morphological (or lexical) form is legitimate for GALLURA. "Let the legs be strong!" (*Techezáqna ha-ragláyim!*) is a modification of "Let the hands be strong!" (*Techezáqna ha-yadáyim!*), the title of a famous labour song by Bialik. Such a temporal flashforward for a story set at the times of the Tower of Babel is a funny transgression (rather than an insipient inconsistency).

"Whatever you shall put under your teeth, you shall find in your legs!" (*Má še-tittnú táchat šinneikhém, timtse'ú be-ragleikhúm!*). This is a Hebrew adaptation of an Aramaic early rabbinic proverb. "Be chewing" (*hevú lo'asín,* associated by PSM with *Laos*), "as for this you were created!" (*še-*

la-zé notsártem! This is evocative of le-khakh notsarta, "for that purpose [of studying] thou hast been created", in *Maxims of the Fathers,* 2:8).

"Every population" (lit.: population population, *'okhlosín 'okhlosín*), "all of them are chewers!" (*kullam la'osot!*). "As they were hearing them saying so" (*Keván še-hayú som'ín 'otám 'omrím ken,* partly a quotation of how the crowd in the Temple used to respond to a given utterance of the High Priest on the Day of Atonement, a day of fast), "their saliva flowed, they paid the price, would take and eat" (*záv hayá rirám, notnín mamón, notlín ve-'okhlín,* with typical Mishnaic wording).

Clearly, obtaining from GALLURA output such as this story from our model literary corpus would be as ambitious a goal as it can get. Anything in the middle would be nice to achieve. See Nissan's (2011b) 150-page discussion of playful narrative explanations.

6 WHENCE AND WITHER? *VON* IN GERMAN ONOMASTICS, AND ELDAD & MEDAD

It is important to realise that knowledge discovery or information extraction as involved in accessing the historical textual canon as well as ontologies and representations of commonsense or encyclopedic knowledge, can be easier in one direction, while very difficult in the other. We exemplify this with an item from our model corpus. Edom (medieval for 'Europeans') in the Land of Ashkenaz (medieval for 'Germany') the text relates, for many generations were eager to insert *fon* (i.e., *von*) before their family names, as it would signal their patrician ancestry.

"What did cause that? The episode of Eldad and Medad caused that", at *Numbers* 11:26–29. Moses appointed seventy elders, but those two did not come, and prophesised nevertheless. Joshua tells Moses to put them under arrest, but Moses retorts: "Are you jealous on my behalf? If only" all the people were prophets.

U-mí yittén (lit.: "And who would give") was rendered, in the canonical Jewish Aramaic translation (the *Targum* by Onqelos) as: *Ra'ena fon,* i.e., "I wish fon", where *fon* (a grammaticalised denominal conjunction) means any of 'face', 'turn', 'that would', or 'lest'. If you were reading Onqelos, you may happen to notice this *locus* serendipitously. But had you begun with eagerness for ennoblement, it would be very difficult to devise an appropriate search that would retrieve a biblical "I want *fon*".

7 CONCLUDING REMARKS

GALLURA is an ambitious project, now in the design phase, requiring the interplay of various agents or coalitions of agents specialised per domain of expertise. Several of these agents have retrieval-intensive requirements. GALLURA has to devise playful etymologies with a backup story to go with. It builds upon the experience and part of the architectural features of DARSHAN — especially how the pool of devices is organised, and the approach to retrieval, which is mostly from the same textual corpora. GALLURA is much more difficult to achieve, but at the stage reached by a number of domains within AI, NLP, IR, and KD, it is in principle feasible. Any progress on any part of the architecture would by itself be a valuable achievement. A global advantage already at present, in this project, is that thanks to manual analysis of many items in the creative writing corpus which is our main model, it is possible to model algorithmically all devices required.

REFERENCES

Baldinger, K. 1973. À propos de l'influence de la langue sur la pensée: Étymologie populaire et changement sémantique parallèle. *Revue de Linguistique Romane,* 37, pp. 241–273.

Bex, F. 2011. *Arguments, Stories and Criminal Evidence: A Formal Hybrid Theory*, Law and Philosophy Series, vol. 92. Springer, Dordrecht.

Braude, W. G. 1982. Midrash as deep peshat. In: S. R. Brunswick (ed.), *Studies in Judaica, Karaitica and Islamica (Presented to Leon Nemoy on his Eighties Birthday).* Bar-Ilan University Press, Ramat-Gan, Israel, pp. 31–38 [English].

Choueka, Y. 1989a. RESPONSA: An operational full-text retrieval system with linguistic components for large corpora. In: E.I. Cuomo (ed.), *Law in Multicultural Societies, Proceedings of IALL, the International Association of Law Libraries Meeting,* Jerusalem, 1985. The Hebrew University, Jerusalem, 1989, pp. 47–82.

Choueka, Y. 1989b. Responsa: A full-text retrieval system with linguistic processing for a 65 million-word corpus of Jewish heritage in Hebrew. In a Special Issue on non-English Interfaces to Databases*, IEEE Data Engineering,* 12(4), pp. 22–31.

Choueka, Y., Cohen, M., Dueck, J., Fraenkel, A.S., Slae, M. 1971. Full-text Document Retrieval: Hebrew Legal Texts (Report on the first phase of the Responsa Retrieval Project). In: M. Minker, S. Rosenfeld (eds.), *Proceedings of the ACM Symposium on Information Storage and Retrieval,* Maryland*,* 1971. Association for Computing Machinery, New York, 1971, 61–79.

Choueka, Y., Fraenkel, A. S., Klein, S.T., Segal, E. 1987. Improved techniques for processing queries in full-text systems. In: C.T. Yu, C.J. van Rijsbergen (eds.), *Proceedings of the Tenth Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval,* New Orleans 1987. ACM, New York, 1987, pp. 306–315.

Coates, R. 1994. Folk etymology. In: R.E. Asher (ed.), *The Encyclopedia of Language and Linguistics,* Pergamon Press, Oxford, Vol. 3, pp. 1267–1270.

Fishbane, M., ed. 1993. *The Midrashic Imagination,* University of New York Press, New York.

HaCohen-Kerner, Y., Mughaz, D. 2010. Estimating the birth and death years of authors of undated documents using undated citations. *Proceedings of the Seventh International Conference on Natural Language Processing (IceTAL 2010),* August 16–18, 2010, Reykjavik, Iceland (LNCS 6233), pp. 138–149. Springer-Verlag, Berlin.

HaCohen-Kerner, Y., Avigezer, T.S.-T., Ivgi, H. 2007. The Computerized Preacher: A prototype of an automatic system that creates a short rabbinic homily [Hebrew]. *B.D.D. (Bekhol Derakhekha Daehu): Journal of Torah and Scholarship* (Bar-Ilan University, Ramat-Gan) 18, pp. 23–46.

HaCohen-Kerner, Y., Beck, H., Yehudai, E., Rosenstein, M., Mughaz, D. 2010a. Cuisine: Classification using stylistic feature sets and/or name-based feature sets. *Journal of the American Society for Information Science and Technology,* 61(8), pp. 1644–1657.

HaCohen-Kerner, Y., Kass, A., Peretz, A. 2010b. A Hebrew Aramaic abbreviation disambiguation system. *Journal of the American Society for Information Science and Technology,* 61(9), pp. 1923–1932.

HaCohen-Kerner, Y., Schweitzer, N., Shoham, Y. 2010c. Automatic identification of biblical quotations in Hebrew-Aramaic documents. *Int. Conf. on Knowledge Discovery and Information Retrieval (KDIR),* pp. 320–325, Oct. 2010, Valencia.

Hartman, G. H., Budick, S., eds. 1986. *Midrash and Literature,* Yale University Press, New Haven, CT.

Hirshman, M. 2006 . Aggadic midrash. Ch. 2 in: S. Safrai, Z. Safrai, J. Schwartz, P. J. Tomson (eds.), *The Literature of the Sages, Second Part,* Royal Van Gorcum, Assen, Netherlands, and Augsburg Fortress Press, Minneapolis, MN, pp. 107–132.

Kirwin, W. 1985. Folk etymology: Remarks on linguistic solving and who does it. *Lore and Language* (Sheffield, U.K.), 4(1), pp. 18–24.

Liu, H., Singh, P. 2002. MAKEBELIEVE: Using commonsense knowledge to generate stories. In *Proc. of the 18th National Conf. on Artificial Intelligence and 14th Conf. on Innovative Applications of Artificial Intelligence,* pp. 957–958.

Lönneker, B., Meister, J. C., Gervás, P., Peinado, F., Mateas, M. 2005. Story generators: Models and approaches for the generation of literary artefacts. In the *ACH/ALLC-2005 Conference Abstracts,* Victoria, BC, Canada, June 15–18, 2005, pp. 126–133.

Nissan, E. 2008. Chance vs. causality, and a taxonomy of explanations. In: M. Negrotti (ed.), *Natural Chance, Artificial Chance,* thematic volume of *Yearbook of the Artificial,* Vol. 5. Peter Lang, Basel, pp. 195–258.

Nissan, E. 2011a. *Computer Applications for Handling Legal Evidence, Police Investigation, and Case Argumentation.* Springer, Dordrecht.

Nissan, E. 2011b. A Study of Humorous Explanatory Tales. In: N. Dershowitz and E. Nissan (ed.), *Language, Culture, Computation: Essays in Honour of Yaacov Choueka*. Springer, Berlin, in press.

Nissan, E., Weiss, H. 1994. The *HyperJoseph* project (2 parts). In: F. Poswick (ed.), *Proc. 4th International Conference on Bible and Computers (AIBI'94),* Amsterdam, August 15–18, 1994. Champion-Slatkine, Geneva & Paris, 1995, pp. 154–162 & 163–173.

Peinado, F., Gervás, P. 2006. Evaluation of automatic generation of basic stories. In a special issue on Computational Creativity, *New Generation Computing*, 24(3), pp. 289–302.

Reeves, J. 1991. *Computational Morality: A Process Model of Belief Conflict and Resolution for Story Understanding,* Tech. Rep. 910017, Comp. Science Dept.. Univ. of California, Los Angeles. ftp: // ftp. cs.ucla.edu/tech-report/1991-reports/910017.pdf

Ritchie, G. 2004. *The Linguistic Analysis of Jokes,* Routledge, London.

Schank, R. C., ed. 1986. *Explanation Patterns: Understanding Mechanically and Creatively,* Lawrence Erlbaum Associates, Hillsdale, NJ.

Schank, R. C., Kass, A., Riesbeck, C. K., eds. 1994. *Inside Case-Based Explanation*, Erlbaum, Hillsdale, NJ.

Stock, O., Strapparava, C. 2005. The act of creating humorous acronyms. *Applied Artificial Intelligence,* 19(2), pp. 131–151.

Stock, O., Strapparava, C., Nijholt, A., eds. 2002. *The April Fools' Day Workshop on Computational Humour: Proceedings of the 20th Twente Workshop on Language Technology (TWLT20),* Trento, Italy, April 2002. University of Twente, The Netherlands.

Strapparava, C., Valitutti, A. 2004. WordNet-Affect: An affective extension of WordNet. *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC 2004),* Lisbon, May 2004, pp. 1083–1086.

Waller, A., Black, R., O'Mara, D. A., Pain, H., Ritchie, G., Manurung, R. 2009. Evaluating the STANDUP pun generating software with children with cerebral palsy. *ACM Transactions on Accessible Computing (TACCESS),* 1(3), article no. 16, at the ACM site.

Walton, D. N. 2004. *Abductive Reasoning,* University of Alabama Press, Tuscaloosa, Alabama.

Zuckermann, G. 2000. *Camouflaged Borrowing: Folk-Etymological Nativization in the Service of Puristic Language Engineering,* D.Phil. Dissertation in Modern Languages, University of Oxford, Oxford.

Zuckermann, G. 2006. "Ety*myth*ological othering" and the power of "lexical engineering". Ch. 16 in T. Omoniyi, J.A. Fishman (eds.), *Explorations in the Sociology of Language and Religion,* Benjamins, Amsterdam, pp. 237–258.